Universidad Autónoma de Madrid
Escuela politécnica superior

Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación

Trabajo Fin de Grado

VERIFICACIÓN DE FIRMA MANUSCRITA ESTÁTICA MEDIANTE REDES NEURONALES CONVOLUCIONALES

Autor: Gustavo Manzano Ramírez
Tutor: Rubén Vera Rodríguez
Ponente: Julián Fiérrez Aguilar

JUNIO 2019

VERIFICACIÓN DE FIRMA MANUSCRITA ESTÁTICA MEDIANTE REDES NEURONALES CONVOLUCIONALES

Autor: Gustavo Manzano Ramírez
Tutor: Rubén Vera Rodríguez
Ponente: Julián Fiérrez Aguilar

Biometrics and Data Pattern Analytics - BiDA Lab
Dpto. de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
JUNIO 2019

Resumen

Resumen

En este trabajo se estudia, desarrolla y evalúa un sistema de reconocimiento biométrico de firma estática sintética gracias a Redes Neuronales Convolucionales. Al principio del trabajo se encuentra un estudio del estado del arte en reconocimiento biométrico haciendo especial énfasis en los sistemas basados en Redes Neuronales Convolucionales.

Una vez entendida la parte más teórica del sistema, se explica en detalle cómo se implementa. Éste extrae características útiles de las imágenes gracias a una Red Neuronal Convolucional entrenada con una base de datos amplia de firma manuscrita estática sintética. Después, a través de un clasificador dependiente del usuario SVM (Support Vector Machines) se halla un hiperplano capaz de separar las características del usuario de las del resto, y de este modo identificar y verificar la identidad del usuario. La Red Neuronal Convolucional crea un modelo único que permite al sistema generalizar a diferentes bases de datos, llegando incluso a mejorar los resultados del estado del arte analizado para firma estática original.

En este trabajo se ha investigado en profundidad la extracción de la firma estática a partir de la dinámica ya que tiene vital importancia para la disposición del sistema, realizando cambios en los parámetros de la firma (tamaño del píxel, presión, etc.). A su vez, se ha investigado el entrenamiento de los clasificadores y como puede afectar variar el número de muestras positivas y negativas. Por lo tanto, demostrando lo crucial que son ciertos parámetros para el correcto funcionamiento del sistema.

La parte experimental se lleva a cabo de dos partes. Primero, se compara con el estado del arte de firma estática para poder decidir el mejor modo de extraer la firma sintética. Y segundo, se utiliza una base de datos de gran extensión para crear un modelo global capaz de generalizar entre distintas bases de datos.

Finalmente, se tratan las conclusiones extraídas, así como las posibles implementaciones y líneas de trabajo que se podrían realizar de cara a mejorar y perfeccionar el sistema en un futuro.

Palabras Clave

Sistema biométrico, CNN, SVM, firma off-line sintética, firma genuina, skilled forgery, random forgery, EER.

Abstract

In this work, a biometric recognition system of synthetic offline signatures is studied, developed and evaluated thanks to Convolutional Neural Networks. At first, there is a study of the state of the art in biometric recognition, with special emphasis on systems based on CNNs.

Once understood the most theoretical part of the system, it is explained in detail how is implemented. It extracts useful characteristics of the images through a Neuronal Convolutional Network. The CNN is trained with a large database of off-line synthetic signatures. Then, using a writer-dependent classifier SVM (Support Vector Machines) a hyperplane is created, capable of separating the characteristics extracted before for each user. Thereby, able to verify and identify the user. The Convolutional Neural Network creates a unique model that allows the system to generalize to different databases, improving the results of the state of the art analyzed for original offline signatures.

The extraction of the off-line signature from the on-line has been investigated in depth since it has a vital importance for the good behavior of the system, performing changes in the signature's parameters such as pixel size, pressure, pen-ups, etc. At the same time, it has been explored how can affect to vary the number of positive and negative samples in the training phase of the classifiers. Therefore, showing how crucial certain parameters are for the proper operation of the system.

The experimental part is carried out in two parts. First, comparing with the state of art of online signature to be able to extract the synthetic signature the best way possible. And second, using a database of great extension to train a model capable of generalizing to different databases.

Finally, it is shown the conclusions extracted, as well as the possible implementations and the lines of work that can be carried out in order to improve and perfect the system in the future.

Key words

Biometric system, CNN, SVM, synthetic offline signatures, genuine signature, skilled forgery, random forgery, EER.

Agradecimientos

En primer lugar, me gustaría agradecer a mi tutor Rubén Vera por brindarme la oportunidad de realizar este trabajo bajo su supervisión, ser tan atento y entregado conmigo y guiarme de manera efectiva ante todas las complicaciones que se han ido interponiendo. También a Rubén Tolosa, quién pese a no ser mi tutor ha estado durante todo el trabajo aportando ideas y mejoras. Agradecer de manera especial a Robert Sabourin por resolvernos dudas acerca de su trabajo que han permitido mejorar el estado del arte. Y por último, a todo el grupo de investigación BiDa Lab por su confianza.

Ahora, dar las gracias a las personas que han estado apoyándome de manera más cercana. Mi familia, tanto mis padres y mi hermana como todos mis tíos y abuelos que se han preocupado durante toda la carrera por mí. Mi grupo de amig@s más cercanos que saben lo importantes que son y el aprecio que les tengo, capaces que consiga despejarme en los momentos más complicados. También a la persona especial que ha aguantado todos mis agobios y malos ratos, te llevas la mejor parte de mí.

Por último, agradecer a mis compañeros de clase, con los que he compartido miles de momentos durante toda la carrera y saben el trabajo que ha supuesto seguir adelante y no rendirse en ninguna circunstancia. Han sido unos años muy bonitos que recordaré con mucha nostalgia. Muchas gracias a todos de corazón.

Gustavo Manzano Ramírez

Junio 2019

Índice general

Índice de Figuras vii

Índice de Tablas viii

1. Introducción 1

1.1. Motivación . . . 1

1.2. Objetivos . . . 2

1.3. Metodología y plan de trabajo . . . 2

2. Estado del arte 4

2.1. Importancia de la biometría en la actualidad . . . 4

2.2. Características y funcionamiento de los sistemas biométricos . . . 5

2.3. Redes Neuronales . . . 7

2.4. Sistemas basados en firma manuscrita . . . 9

2.4.1. Introducción . . . 9

2.4.2. Sistemas tradicionales . . . 10

2.4.3. Redes Neuronales Convolucionales Profundas . . . 12

2.4.4. Extracción de características para verificación de firmas a través de Redes Neuronales Convolucionales . . . 13

3. Sistema propuesto 18

3.1. Introducción . . . 18

3.2. Entrenamiento de la CNN . . . 18

3.3. Entrenamiento de los clasificadores Writer-Dependent . . . 19

4. Bases de Datos 21

4.1. Introducción . . . 21

4.2. DeepSignDB . . . 21

4.3. Preprocesado de los datos de firma on-line . . . 23

4.4. Preprocesado de las imágenes de firma off-line . . . 24

4.5. Base de Datos definitiva . . . 25

4.6. División de la Base de Datos . . . 25

5. Desarrollo experimental **27**

5.1. Creación de la arquitectura y adaptación al entorno 27

5.2. Obtención de la mejor versión de firma off-line sintética 27

5.2.1. Protocolo experimental 27

5.2.2. Desarrollo experimental 28

5.2.3. Evaluación de los resultados 30

5.3. Creación de un modelo global 31

5.3.1. Protocolo experimental 31

5.3.2. Desarrollo experimental 33

5.3.3. Evaluación de los resultados 33

5.4. Experimentación con los clasificadores 34

5.4.1. Protocolo y desarrollo experimental 34

5.4.2. Evaluación de los resultados 35

5.5. Comparación con el estado del arte 35

6. Conclusiones y trabajo futuro **36**

Glosario de acrónimos **37**

Bibliografía **38**

Índice de Figuras

1.1. Diagrama del plan de trabajo seguido. 2

2.1. Esquema del funcionamiento de un sistema de reconocimiento biométrico de firma. Etapa de registro/Fase de entrenamiento. Figura adaptada de [4]. 6

2.2. Esquema del funcionamiento de un sistema de reconocimiento biométrico de firma. Modo de Identificación. Figura adaptada de [4]. 7

2.3. Esquema del funcionamiento de un sistema de reconocimiento biométrico. Modo de Verificación. Figura adaptada de [4]. 7

2.4. Arquitectura típica de un sistema de verificación de firma. 10

2.5. Ilustración de la estructura de la CNN usada. Fuente: [1] 13

2.6. La base de datos GPDS es separada en un conjunto de explotación y un conjunto de desarrollo. El conjunto de desarrollo se usa para el aprendizaje de las características y hacer todas las decisiones del modelo. El conjunto de explotación representa los nuevos usuarios que llegan al sistema, en donde se entrenan los clasificadores WD solo usando firmas genuinas. Fuente: [1] 14

4.1. Dispositivos de captura utilizados en e-BioSign DS1. Fuente: [18]. 23

4.2. Elementos de la firma on-line (izquierda) y su correspondiente firma off-line(derecha). Fuente: [2] 24

5.1. Comparación de la firma sintética con la firma original. De izquierda a derecha: versión 1, versión 2, versión 3 y versión 4. 29

5.2. Comparación de la firma sintética con la firma original. Izquierda, versión 5. Derecha, versión 6 29

5.3. Comparación de utilizar diferentes tamaños de pixel en la extracción de la firma sintética con modelos específicos (versiones de 1-4) frente a utilizar el modelo SigNet. 31

5.4. Comportamiento de las diferentes BBDD. Línea continua, utilizando el 100 % del conjunto de desarrollo. Línea discontinua, utilizando el 50 %. 34

Índice de Tablas

2.1. Tipos de rasgos biométricos. 5

2.2. Comparación cualitativa de varios rasgos biométricos. A=Alto, M=Medio, B=Bajo. 6

2.3. Resumen de las capas de la CNN. Fuente: [1] 15

2.4. Lista de las operaciones de prealimentación. Fuente: [1] 15

2.5. Comparación con el estado del arte en MCYT. EER de skilled forgeries. Fuente: [1] 17

4.1. Características de las bases de datos utilizadas. SG=Dispositivo Samsung, W=Tablet Wacom, st=stylus, d=dedo. Fuente: [2] 22

4.2. Distribución de la Base de Datos en número de firmas por usuario. Genuinas y falsificaciones. 25

4.3. Distribución de la Base de Datos en el Conjunto de Desarrollo y Explotación . . 26

5.1. División de la base de datos MCYT. 28

5.2. División del SVM para cada usuario de MCYT-72. 28

5.3. Tabla de resultados para MCYT-72 sintética con diferentes versiones y comparando el uso del modelo SigNet con el específico de esa versión. 30

5.4. Tabla de resultados para MCYT-72 sintética con diferentes versiones y comparando el uso del modelo SigNet con el específico de esa versión. 31

5.5. Número de firmas utilizadas para el entrenamiento de la CNN. 32

5.6. Protocolo experimental de los clasificadores SVM para los diferentes experimentos según cada base de datos 32

5.7. Resultados para los diferentes experimentos según las bases de datos. 33

5.8. Número de usuarios y del conjunto que provienen introducidos como muestras negativas en el clasificador y su correspondiente EER, tanto Skilled (user thresholds) como Random (users thresholds). Características extraídas con el modelo global. 35

5.9. Comparación con el estado del arte en MCYT. EER de skilled forgeries. 35

1 Introducción

1.1. Motivación

Debido al gran avance tecnológico de los últimos tiempos, se ha expandido masivamente el uso de dispositivos móviles para realizar todo tipo de tareas, desde pagos on-line a gestiones administrativas. Por lo que ha nacido la necesidad de desarrollar nuevos métodos de autentificación, más amigables, que suplan las carencias de los sistemas tradicionales en donde los usuarios conocen PINs o contraseñas que pueden ser fácilmente hackeados. Esta solución recae en los sistemas de autentificación basados en la información biométrica.

La autentificación biométrica es el proceso de verificar la identidad de un individuo a través de sus características físicas o conductuales. Desde el uso de la huella dactilar hasta el reconocimiento facial o de voz. Cada día estos sistemas están más aceptados por la sociedad y se implementan en todo tipo de sistemas de seguridad. La firma manuscrita siempre ha sido el método tradicional utilizado para la verificación del usuario y autentificación de documentos legales debido a su fácil implementación. La gran expansión de dispositivos táctiles de los últimos años ha posibilitado el crecimiento de los sistemas de reconocimiento biométrico basados en la escritura y la firma manuscrita.

Por otro lado, el reconocimiento biométrico basado en firma es un problema de alta complejidad. Existe una variabilidad tanto en el dispositivo de captura (e.g. smartphones, tablets, etc.) como en el útil de escritura (e.g. stylus y dedo), y a su vez una alta variabilidad entre firmas de un mismo usuario (intra-clase) y una baja entre firmas de diferentes usuarios (inter-clase). De todos modos, el funcionamiento de los sistemas basados en redes neuronales profundas, que tienen un alto rendimiento cuando la base de datos es amplia, permiten poder desarrollar sistemas robustos que extraigan características relevantes independientemente del escenario de captura.

Por lo tanto, la principal motivación es desarrollar un sistema de verificación de firma manuscrita estática que generalice correctamente para diferentes bases de datos, es decir, que tenga robustez ante cualquier escenario, ya sea interoperabilidad de dispositivos o variabilidad de usuarios.

Cabe destacar que la firma estática (off-line) carece de información temporal por lo que cuenta con una gran desventaja respecto a la firma dinámica (on-line). Pero debido al uso de redes neuronales convolucionales (CNN), entrenadas para extraer características relevantes de imágenes, puede llegar a competir con la firma dinámica.

1.2. Objetivos

Este trabajo se ha llevado a cabo con el propósito de cumplir los siguientes objetivos:

1. Creación de un sistema de verificación de firma estática gracias a los trabajos previos del estado del arte que hacen uso de redes neuronales convolucionales [1]. Es decir, compresión del código que permita la realización de experimentos y estudio del arte de las arquitecturas utilizadas en trabajos previos.

2. Investigación de la adaptación de la firma **off-line sintética** (extraída a partir de la firma on-line) al sistema, en vez del uso de firma off-line original. Se han tenido en cuenta parámetros de la firma como puede ser el grosor, la presión o la información presente en los trazos de vuelo (pen-ups) que normalmente no se considera en la firma off-line.

3. Adecuación del sistema, así como exploración de los parámetros óptimos para los escenarios descritos anteriormente (e.g. multi-sesión, multidispositivo, etc.).

4. Evaluación del sistema de verificación para la base de datos **DeepSignDB (DeepSign Database) off-line sintética**. [2]

1.3. Metodología y plan de trabajo

Para alcanzar los objetivos establecidos en este Trabajo Fin de Grado, se ha seguido la metodología que se muestra en la Fig. 1.1, detallada a continuación.

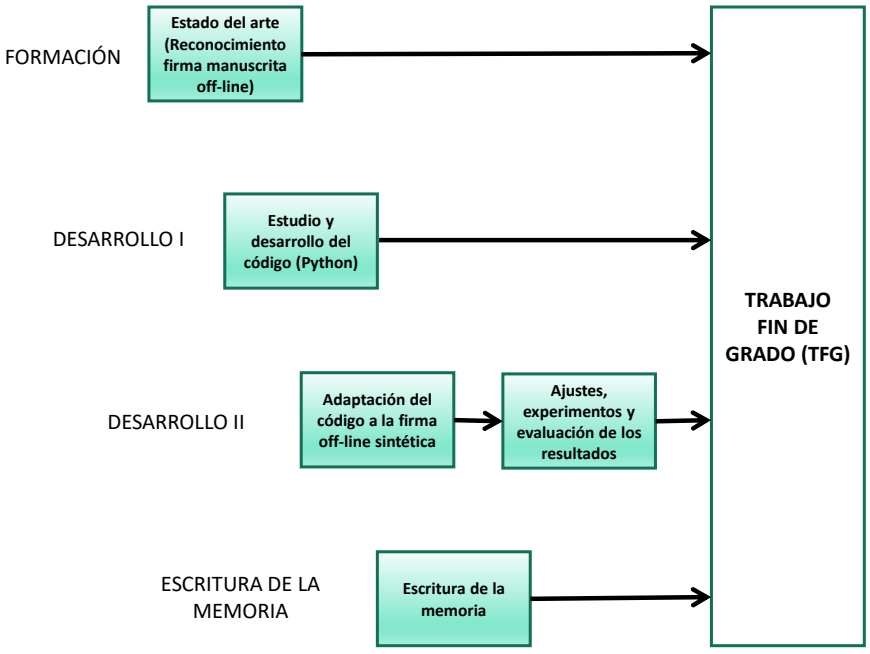

Figura 1.1: Diagrama del plan de trabajo seguido.

- **Estudio del estado del arte.** Lo más importante a la hora de llevar a cabo un proyecto es obtener la información y los conocimientos necesarios para realizarlo. La formación se

ha basado en la lectura y estudio de las características básicas del reconocimiento biométrico, profundizando posteriormente en el estado del arte de la firma manuscrita estática. Para ello se han utilizado libros, publicaciones y cursos que profundizan en el aprendizaje automático y más concretamente en la verificación de firma manuscrita a través de redes neuronales convolucionales. De este modo, poder conocer las tecnologías más punteras y que mejor se pueden adaptar a este trabajo.

- **Estudio y desarrollo del código.** Es necesario realizar un proceso de familiarización con las herramientas y programas disponibles. En este caso, se trata de la comprensión del software utilizado (e.g. Lasagna, Keras, Torch, TensorFlow).
- **Adaptación del código y evaluación.** Una vez obtenidos los conocimientos necesarios y haberse familiarizado con el entorno de programación es necesario hacer una adaptación de la base de datos, ya que normalmente los sistemas de verificación usan las firmas off-line originales, y en este caso se trata con firma off-line sintética. Esto permite que se pueda hacer la experimentación, así como la evaluación del sistema.
- **Escritura de la memoria.** Finalmente, tras llevar a cabo los diferentes experimentos, analizar y evaluar los resultados obtenidos, se procede a realizar un estudio y una comparación con los resultados del estado del arte para desarrollar la memoria del presente Trabajo Fin de Grado.

2

Estado del arte

2.1. Importancia de la biometría en la actualidad

Debido a la expansión tecnológica de los últimos años, en la actualidad existe una gran globalización e interconectividad. Es por esto que los sistemas de verificación de usuarios se están volviendo más complejos, exigiendo una mayor fiabilidad. La biometría se usa en muchos ámbitos sociales entre los que destacan los siguientes:

- **Dispositivos móviles**: En los últimos cinco años ha predominado la implantación en los teléfonos móviles de lectores de huella dactilar para hacer el desbloqueo del dispositivo. Incluso los dispositivos más recientes incluyen desbloqueo facial, es decir, a través del reconocimiento facial. De este modo, se sustituyen los métodos más tradicionales como el PIN o el patrón de desbloqueo, que son más vulnerables a hackeos. Recientemente se han llevado a cabo estudios sobre la autentificación del usuario gracias a la interacción con la pantalla. [3] [4].

- **Ámbito bancario**: Cada vez más compañías bancarías permiten confirmar pagos o transacciones desde el propio dispositivo gracias a los lectores de huella dactilar. Se está realizando una fuerte inversión en la seguridad de los sistemas para evitar fraudes y hackeos que puedan suponer pérdidas incalculables. Los sistemas de verificación de firma manuscrita están ganando gran importancia ya que son los sistemas más extendidos en todo el mundo para validación de documentos.

- **Ámbito turístico**: Algunos países ya se plantean usar el reconocimiento biométrico para controlar la inmigración en los aeropuertos. Por ejemplo, China planea implementar en los Juegos Olímpicos de 2020 reconocimiento facial para todos los oficiales olímpicos y representantes de los medios. Tendrán sus rostros escaneados y relacionados con sus identificaciones, antes de que se les permita el ingreso a los lugares, de este modo esperan que mejore la seguridad y evitar actos terroristas[5].

- **Ámbito forense**: Desde hace muchos años el reconocimiento de criminales gracias al uso de la huella dactilar ha sido muy importante en los procesos jurisdiccionales y en el ámbito criminalístico[6].

2.2. Características y funcionamiento de los sistemas biométricos

El reconocimiento biométrico se basa en identificar patrones únicos para cada usuario en campos muy variados. Estos patrones son extraídos de ciertos rasgos biométricos, que se pueden clasificar en rasgos **físicos** (e.g. la huella dactilar) o rasgos **de comportamiento** (e.g. la firma), ver Tabla 2.1. La principal ventaja de la biometría es su gran dificultad para ser olvidada, robada o perdida.

Tipos de rasgos biométricos	
Físicos	**De Comportamiento**
Cara	Escritura
Geometría de la mano	Modo de andar
Iris	Gestos
Venas de retina	Firma manuscrita
Voz	Voz

Cuadro 2.1: Tipos de rasgos biométricos.

Para poder considerar los rasgos mencionados antes como biométricos, han de cumplir con los siguientes requisitos[7]:

- *Universalidad*: característica que tienen que tener todas las personas.
- *Unicidad*: se tiene que poder diferenciar entre dos personas cualquiera en los que al rasgo biométrico se refiere.
- *Permanencia*: la característica debe ser parcialmente invariante durante un periodo de tiempo.
- *Mensurabilidad*: el rasgo debe ser cuantificable.

A su vez, hay que considerar otra serie de características para poder denominar a un sistema biométrico como práctico. Éstas son las siguientes:

- *Rendimiento*: debe tener una baja tasa de error.
- *Aceptabilidad*: la sociedad debe estar familiarizada con el sistema para su sencillo desarrollo.
- *Evitabilidad*: debe tener cierta fiabilidad frente a métodos fraudulentos o falsificaciones.

En la Tabla.2.2 se aprecian las características nombradas anteriormente en diferentes rasgos biométricos. Se puede ver que ninguno de los rasgos cumple todas ellas. Cada uno cuenta con sus ventajas y desventajas.

Dentro del funcionamiento de los sistemas biométricos, previamente hay una etapa de **registro**, ver Fig. 2.1. En donde se recopila la información del usuario y se almacena en la base de datos a través de sensores capaces de medir el rasgo biométrico.

Una vez hecho el registro del usuario, los sistemas biométricos se componen de dos modos, o ambos conjuntamente, dependiendo del contexto en el que se vaya a realizar la aplicación.

Rasgo Biométrico	Universalidad	Unicidad	Permanencia	Mensurabilidad	Rendimiento	Aceptabilidad	Evitabilidad
ADN	A	A	A	B	A	B	B
Oreja	M	M	A	M	M	A	M
Cara	A	B	M	A	B	A	A
Termograma facial	A	A	B	A	M	A	B
Venas de la mano	M	M	M	M	M	M	B
Huella dactilar	M	A	A	M	A	M	M
Forma de andar	M	B	B	A	B	A	M
Geometría de la mano	M	M	M	A	M	M	M
Iris	A	A	A	M	A	B	B
Huella palmar	M	A	A	M	A	M	M
Olor	A	A	A	B	B	M	B
Retina	A	A	M	B	A	B	B
Firma	B	B	B	A	B	A	A
Forma de teclear	B	B	B	M	B	M	M
Voz	M	B	B	M	B	A	A
Escritura	B	B	B	A	B	A	A

Cuadro 2.2: Comparación cualitativa de varios rasgos biométricos. A=Alto, M=Medio, B=Bajo.

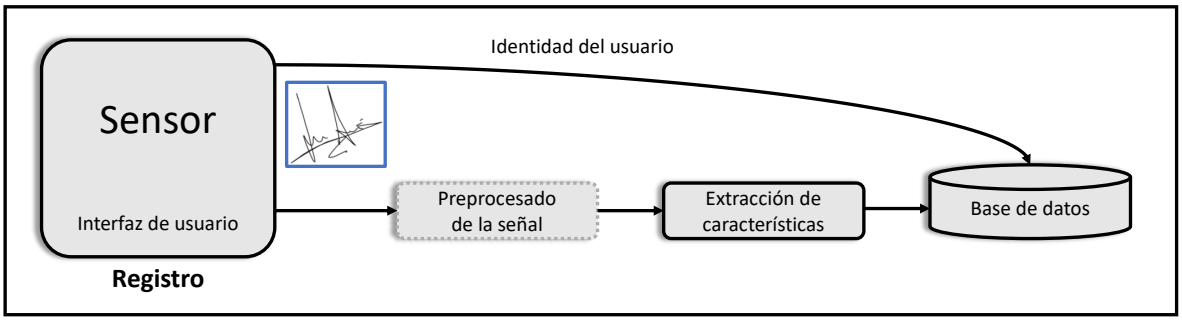

Figura 2.1: Esquema del funcionamiento de un sistema de reconocimiento biométrico de firma. Etapa de registro/Fase de entrenamiento. Figura adaptada de [4].

1. **Sistemas de identificación**: el sistema valida la identidad de un usuario comparando sus rasgos con los datos almacenados en la base de datos. Es una comparación uno-a-muchos, por lo que supone un coste computacional bastante alto. Este modo proporciona la identidad del usuario como salida, o un mensaje indicando que el usuario no pertenece a dicha base de datos, ver Fig. 2.2.

2. **Sistemas de verificación**: el sistema verifica que las características son del usuario en cuestión, es decir, que no se trata de una falsificación. Compara uno-a-uno con los patrones de ese usuario almacenados en la base de datos. La salida del sistema es binaria: identificación positiva, en el caso de que el sistema detecte cierto nivel de similitud entre los rasgos comparados, o negativa, en el caso contrario, ver Fig. 2.3.

Los sistemas de verificación se basan en dos tipos de errores:

- **Falsa Aceptación (FA)**: error producido cuando un usuario se intenta hacer pasar por otro y el sistema lo reconoce como genuino.
- **Falso Rechazo (FR)**: error producido cuando un usuario genuino es rechazado por el sistema como si fuera un usuario impostor.

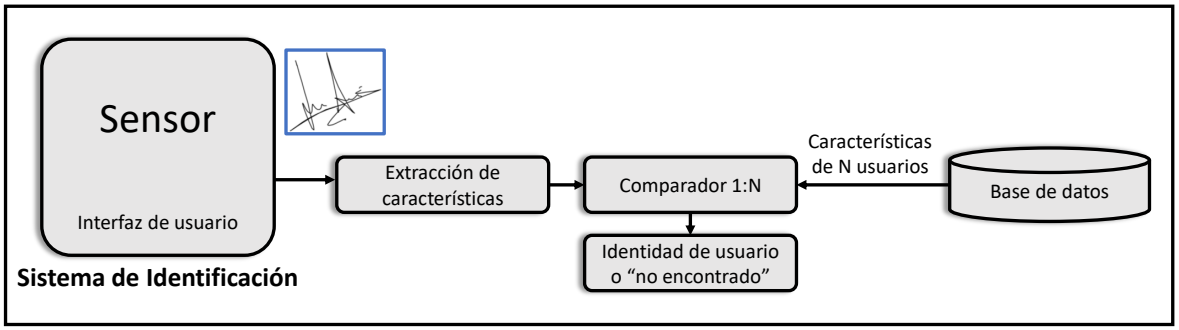

Figura 2.2: Esquema del funcionamiento de un sistema de reconocimiento biométrico de firma. Modo de Identificación. Figura adaptada de [4].

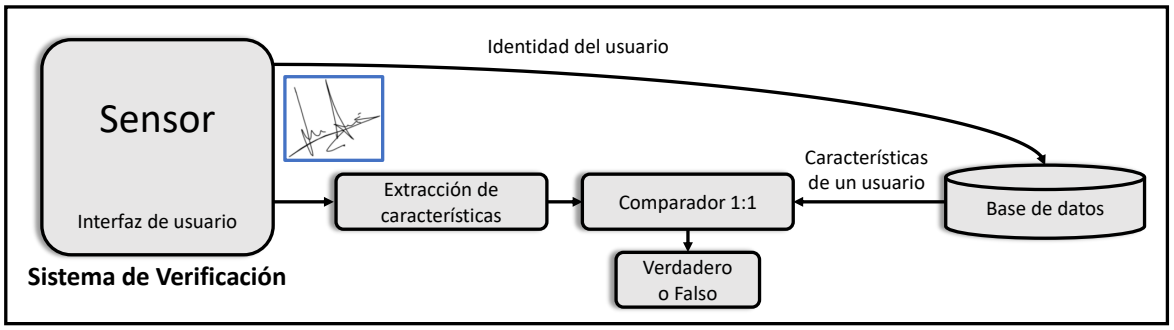

Figura 2.3: Esquema del funcionamiento de un sistema de reconocimiento biométrico. Modo de Verificación. Figura adaptada de [4].

A partir de estos dos errores, del entrenamiento de un sistema de verificación biométrico y de un umbral de decisión, se puede obtener la Tasa de Falsa Aceptación (False Aceptance Rate, FAR) y la Tasa de Falso Rechazo (False Rejection Rate, FRR). Con ellas, aparece un nuevo método de medida del rendimiento del sistema, que aporta más información que las tasas de acierto o de error. Se trata del **Equal Error Rate (EER)**, que consiste en la tasa de error del sistema a un determinado umbral de decisión cuando se cumple la condición FAR = FRR. En este trabajo se va a hacer uso de este método de medida para expresar los resultados a la hora de evaluar el sistema .

Existen dos tipos de falsificaciones en los sistemas de verificación:

- **Skilled Forgery**: cuando el sistema da por positivo las características de un usuario que intencionandamente ha intentado hacerse pasar por otro usuario.
- **Random Forgery**: cuando el sistema da por positivo las características de un usuario cualquiera de la base de datos como otro usuario que no le corresponde.

2.3. Redes Neuronales

Las **Redes Neuronales** (*Neural Networks, NN*) son sistemas formados por varios elementos de procesamiento simples, altamente conectados, que procesan la información de entrada a través de su respuesta de estado dinámico [8] [9]. Están inspiradas en la estructura y funcionamiento del córtex cerebral, pero a menor escala y con conexiones más sencillas. Las redes neuronales tienen tres características fundamentales:

1. **Auto-Organización y Adaptabilidad**: utilizan algoritmos de aprendizaje adaptativo y auto-organización permitiendo un procesado más robusto.
2. **Procesado no Lineal**: permite a la red aumentar su inmunidad frente al ruido y aumentar su capacidad de aproximar funciones o de clasificaciar datos.
3. **Procesado Paralelo**: utiliza un gran número de nodos con un alto nivel de interconectividad.

Las neuronas se suelen denominar *nodos*. Reciben varios valores de entrada (inputs) y producen un valor de salida (output) en función de los pesos asociados a cada input. El valor de los pesos se modifica en el proceso de aprendizaje. Cada neurona aplica una función que suma los valores de entrada ponderados mediante los pesos. También se define el valor umbral o **threshold** a partir del cual se hace la decisión del output.

Las características de los nodos son necesarias para el procesado de imágenes (estático) y de señales temporales (dinámico). La arquitectura generada a partir de la combinación de los nodos combina un procesado adaptativa paralelo con interconexiones jerárquicas. Se divide en dos fases:

- **Fase de entrenamiento**: se utiliza un conjunto de datos para que la red vaya adaptando sus pesos lo mejor posible definiendo un modelo de red neuronal. Los pesos se van calculando de manera iterativa, con el objetivo de minimizar una **función de coste** que mide el error entre la salida real que se desea obtener y la salida de la red neuronal.
- **Fase de test**: se pone a prueba el modelo extraído de la fase de entrenamiento con nuevos datos. Cabe la posibilidad que el modelo se haya adaptado demasiado a las características del conjunto de entrenamiento (i.e. *overfitting*), no siendo capaz de generalizar su aprendizaje a casos nuevos.

El hecho de utilizar datos etiquetados o no etiquetados durante el entrenamiento da lugar a tres tipos de redes, aunque destacan dos de ellas:

- **Redes neuronales de aprendizaje supervisado**: estas redes son las más populares, a priori se conocen las salidas de todos los datos de entrenamiento y evaluación. En general el error es menor en este tipo de redes, aunque existe un alto riesgo de generalización por sobreentrenamiento o sobreajuste.
- **Redes neuronales de aprendizaje no supervisado**: en este caso, el conjunto de datos de entrenamiento solo tiene los patrones de entrada. Por lo tanto, la red se adapta con las experiencias de los patrones de entrenamiento anteriores, intentando minimizar el error entre grupos o clases.

Dentro de la arquitectura de la red neuronal se denomina capa a un conjunto de interconexiones entre los nodos. Estas capas están conectadas por una red de pesos de conexión que pueden ser de 4 tipos: hacia delante, hacia atrás, lateral o de retardo.

Dependiendo del número de capas se pueden ver los siguientes tipos de redes neuronales:

- **Redes monocapa**: los nodos se conectan lateralmente. Algunos de ellos permiten conexiones consigo mismos, dando lugar a redes monocapa autorecurrentes. Las más representativas son la red de Hopfield, la red de memoria asociativa y las máquinas estocásticas de Boltzmann [10] y Cauchy.

- **Redes multicapa**: los nodos están unidos por diferentes tipos de conexiones. Las redes neuronales multicapa suelen tener tres capas: la capa de entrada (*input layer*), la capa oculta (*hidden layer*) y la capa de salida (*output layer*). Lo más común es que las conexiones entre las capas sean hacia delante (conexiones *feedforward*), pero también existen redes en las que puede haber conexiones de capas hacia atrás (conexiones *feedback* o retroalimentadas).

A continuación, se van a definir las redes neuronales más utilizadas hoy en día:

- **Redes Neuronales Convolucionales (CNN)**: son un tipo especializado de redes neuronales que han conseguido un gran éxito en diferentes ámbitos. Su nombre indica que utilizan una operación matemática llamada **convolución** en al menos una de sus capas [8]. Los datos que utiliza pueden ser series temporales, pensados como datos de una dimensión tomados regularmente cada cierto periodo de tiempo, o imágenes, pensadas como una rejilla de dos dimensiones de píxeles. Una de las características más importantes de las CNN es que en cada capa se utiliza una función denominada *Pooling*. Esta función lo que hace es reemplazar la salida de la red en una cierta posición por el resultado de aplicar una operación a los valores vecinos. Las Redes Neuronales Convolucionales son muy importantes en la biometría actual, consiguiendo muy buenos resultados en ámbitos muy diferentes como reconocimiento facial [11], reconocimiento espacial de pasos [12] o reconocimiento de escritura [13].

- **Redes Neuronales Recurrentes (RNN)**: son un tipo específico de arquitectura Deep Learning (DL) para modelar secuencias temporales de datos. Su rango de aplicación es muy variado, abarcando campos desde el reconocimiento de voz [14] a problemas biomédicos [15]. Son modelos con capas ocultas (*hidden layers*) autoconectadas. Un beneficio de este tipo de redes neuronales es que la memoria de las entradas previas se mantiene en el estado interno de la red, permitiendo hacer uso de la información del pasado.

2.4. Sistemas basados en firma manuscrita

2.4.1. Introducción

Como se ha podido observar en el Tabla 2.2, la firma manuscrita es un rasgo biométrico que tiene unas ventajas aprovechables para desarrollar un sistema de verificación. En primer lugar, tiene un alto nivel de aceptabilidad ya que la firma ha sido comúnmente utilizada como método de validación y autenticación de documentos financieros y legales durante muchos años. A su vez, tiene un alto nivel de mensurabilidad, ya que hoy en día existen un gran número de vías para obtener firmas, como documentos firmados de los que se pueden escanear como multitud de dispositivos móviles (e.g. smartphones, tablets).

Pero la problemática de los sistemas de verificación basados en firma manuscrita sigue siendo un reto debido a los tres siguientes puntos:

- **Alta variabilidad intra-clase**: la firma es un rasgo que está influenciado por las condiciones físicas y emocionales en las que se encuentra el firmante. Esto introduce variabilidades intra-clase entre firmas genuinas de un mismo usuario, que debe tenerse en cuenta a la hora de verificar su identidad. Es decir, en un corto período de tiempo puede existir una diferencia notable entre firmas de un mismo usuario.

- **Baja variabilidad inter-clase**: se puede dar el caso, que dos usuarios tengas firmas con una similitud importante. También, se debe tener en cuenta posibles intentos de falsificación, que dan lugar a firmas muy parecidas a las genuinas.

Además, existen dos tipos de sistemas de verificación de firma en función de qué información se utiliza:

- **Sistemas de verificación de firma on-line**: las firmas son capturadas con distintos dispositivos electrónicos (e.g. smartphone, tablets), permitiendo la obtención de la información biométrica del usuario durante todo el proceso de realización, como puede ser la la presión ejercida sobre el dispositivo, la inclinación del bolígrafo, etc.
- **Sistemas de verificación de firma off-line original**: las firmas se hacen con un tintero y la información se digitaliza mediante escáneres ópticos. De este modo se obtienen imágenes digitales de ellas.
- **Sistemas de verificación de firma off-line sintética**: las imágenes digitales de las firmas se obtienen a partir de los datos capturados por los dispositivos de firma on-line. Es decir, se dibuja el trazo digitalmente gracias a los datos de la firma on-line, no un escáner. Por lo tanto, se puede llegar a hacer uso de la información biométrica, que los sistemas de firma off-line originales carecen (presión, información de vuelo, etc.). Éste es el caso tratado en este trabajo.

La complejidad es otro rasgo importante de la firma manuscrita. Éstas se pueden clasificar en firmas de baja, media y alta complejidad. En estudios previos se ha demostrado que los sistemas de verificación de firma son muy sensibles a la complejidad de la firma [16]. Pero este no es el único escenario de estudio posible. Escenarios actuales como interoperabilidad de dispositivos o uso del dedo como útil de escritura suponen un gran desafío para los sistemas de firma manuscrita on-line.

2.4.2. Sistemas tradicionales

Los sistemas de verificación de firma manuscrita dinámica cuentan con una estructura básica. Esta estructura puede ser modificada en función del estudio que se vaya a realizar, pero sus elementos típicos son los que se muestran en la Fig. 2.4, explicados a continuación:

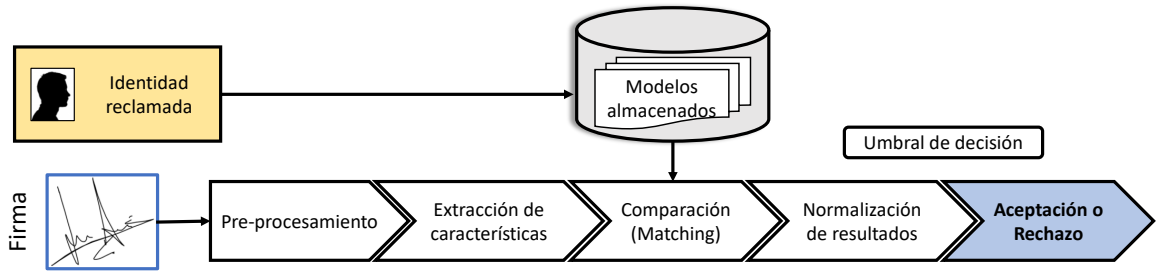

Figura 2.4: Arquitectura típica de un sistema de verificación de firma.

1. **Captura de datos**: los dispositivos de uso genérico (e.g. tablet Samsung con stylus) permiten capturar las coordenadas y la presión de la firma. En cambio, los dispositivos de captura específicos para la adquisición de escritura o firma manuscrita on-line (e.g. tablets

Wacom) pueden capturar también el ángulo de inclinación del bolígrafo durante el periodo de firma y la trayectoria del bolígrafo durante los pen-ups (periodo de tiempo en el que se levanta el útil de escritura de la pantalla entre trazos de la firma). La información de los pen-ups resulta de gran importancia, como se muestra en varios estudios del estado del arte [17] [18]. Las señales se muestrean temporalmente a frecuencias entre 100 y 200Hz, cumpliendo con el criterio de Nyquist, ya que la mayor frecuencia observada a la hora de realizar una firma es de 20-30Hz [19].

2. **Etapa de preprocesamiento**: una vez capturados los datos, suele haber un preprocesamiento para filtrar ruido, diezmar la señal discreta para eliminar muestras repetidas, eliminar ceros iniciales y finales, etc.

3. **Extracción de características**: es el proceso de obtención de características discriminatorias. Deriva en dos sistemas: 1) Sistemas basados en características globales (e.g. número de pen-ups, duración o velocidad de la firma) y 2) Sistemas basados en funciones temporales (e.g. presión, trayectoria). Los algoritmos de selección de características se centran en reducir el tamaño del vector de características para optimizar el rendimiento del sistema para un criterio dado (como por ejemplo EER). Existen muchos tipos de algoritmos de selección de características (i.e. ***Scalar Feature Selection***, ***Sequential Forward/Backward Selection***).

4. **Registro**: existen dos tipos de registros. Un registro basado en un modelo estadístico de cada usuario a partir de un conjunto de firmas genuinas de entrenamiento y un registro basado en referencia donde las características de cada una de las firmas son usadas como plantillas y después se compara la firma entrante con el conjunto de plantillas guardadas para ese usuario.

5. **Cálculo de similitud (*matching*)**: los sistemas basados en características globales suelen usar técnicas de medida de distancias. Por otro lado, los sistemas basados en funciones temporales suelen usar técnicas que comparan modelos de firmas, como los tres siguientes:

   - ***Hidden Markov Models* (HMM)**: técnica muy utilizada en biometría, destacando en reconocimiento de voz, aunque ha sido utilizada también en sistemas de reconocimiento de firma dinámica [20]. Es un proceso doblemente estocástico en el que el sistema a modelar se considera un proceso de Markov de parámetros desconocidos. Estos parámetros se determinan mediante observaciones, que son modeladas por GMMs (*Gaussian Mixture Models*).
   - **Dynamic Time Warping (DTW)**: técnica que permite comparar y encontrar la alineación óptima entre dos secuencias temporales de diferente longitud. Permite encontrar regiones correspondientes entre dos secuencias o estimar la similitud entre ellas. Aunque el algoritmo DTW ha sido de utilidad en muchas disciplinas como el reconocimiento de voz, la extracción de datos, la robótica y la medicina, también se ha utilizado en el campo de verificación de firma on-line [21].
   - **Support Vector Machines (SVM)**: las máquinas de vectores de soporte son un método de aprendizaje para resolver problemas de clasificación y regresión. Una Máquina de vectores de soporte (SVM) es un clasificador discriminativo definido formalmente por un hiperplano separador. Se verán en profundidad en la Sección 3.3. ya que serán usadas en este trabajo.

6. **Normalización de scores**: etapa muy importante. Tras realizar el cálculo de similitud, se suelen normalizar los resultados a un rango común. En [22] se estudian algunas de las técnicas de normalización más utilizadas.

2.4.3. Redes Neuronales Convolucionales Profundas

Las Redes Neurales Convolucionales Profundas (CNN) son redes neuronales multicapa, con una o varias convolucionales, de diferentes tamaños de kernel intercalados por capas de agrupación (pooling), que minimizan la salida de sus convoluciones antes de alimentar a la capa siguiente. En general, se elige una función de pérdidas diferenciable para que se pueda aplicar el descenso de gradiente y se puedan optimizar los pesos de la red. Los pesos de las diferentes capas se actualizan utilizando la técnica comúnmente conocida como backpropagation. Las optimizaciones se organizan por batches (lotes) ya que no se pueden aplicar a todos los datos de entrenamiento y ofrecen una alternativa justa para optimizar la red. Las redes neuronales convolucionales generalmente están compuestas por un conjunto de capas se pueden agrupar por sus funcionalidades [23]:

- **Input Layer**: capa de entrada de la red. Las imágenes tienen que tener las mismas dimensiones que ésta para el funcionamiento de la red.
- **Convolution Layer**: se trata de una convolución 2D de las entradas (imágenes de las firmas). En la convolución, cada píxel de salida es un combinación lineal de los píxeles de entrada. El filtro tiene la misma cantidad de capas que los canales de volumen de entrada, y el volumen de salida tiene la misma profundidad que la cantidad de filtros
- **Activation Layer**: se utiliza para aumentar la no linealidad de la red sin afectar los campos receptivos de las capas de convolución.
- **Pooling Layer**: también denominada capa de reducción, se encargara de reducir la cantidad de parámetros y quedarse con las características más comunes. Generalmente se coloca después de la capa convolucional y afecta a las dimensiones espaciales del volumen y no a las dimensiones de profundidad.
- **Fully-Connected Layer**: las neuronas en una Fully-Connected Layer (capa completamente conectada) tienen conexiones completas a todas las activaciones en la capa anterior. Por lo tanto, sus activaciones pueden calcularse con una multiplicación de matrices seguida de un desplazamiento de sesgo.
- **Softmax**: es un tipo especial de capa de activación y se utiliza generalmente al final de las salidas de la capa Fully-Connected. Se puede ver como un normalizador sofisticado y produce un vector discreto de distribución de probabilidad. Es muy conveniente su uso cuando se combina con la función de pérdidas de entropía cruzada.
- **Regularization**: se utiliza principalmente para evitar el sobreajuste con una gran cantidad de datos de entrenamiento.
- **Dropout**: es una técnica específica de regularización. Consiste en desconectar un porcentaje de las neuronas en cada iteración del entrenamiento, lo que hace que se evite de forma efectiva el sobreajuste al reducir la correlación entre las neuronas.
- **Batch Normalization**: hace que las redes sean robustas para una mala inicialización de pesos en la red. Generalmente se inserta justo antes de las capas de activación. Reduce el cambio de covarianza normalizando y escalando las entradas. Los parámetros de escala y desplazamiento son entrenables para evitar perder estabilidad de la red.
- **Output Layer**: última capa de la red que produce las variables de salida.

2.4.4. Extracción de características para verificación de firmas a través de Redes Neuronales Convolucionales

Introducción

En esta sección se explicará en detalle el sistema de verificación propuesto por la Universidad de Québec y la de Paraná, más concretamente por Luiz G. Hafemann, Robert Sabourin y Luiz S. Oliveira ya que es el sistema que se toma de baseline en este trabajo [1].

El sistema utiliza una CNN como extractor de características y después a través de un SVM específico para cada usuario, es decir, **Writer-Dependent** (WD), determina si se trata de una firma genuina, o una falsificación, obteniendo los mejores resultados para firma estática hasta la fecha para firma estática original.

Se trata de un enfoque en dos fases para el problema: un aprendizaje de características independiente del usuario y una segunda fase de clasificación dependiente del usuario. La idea central es aprovechar los datos de muchos usuarios para crear un espacio de características intrínsecas de firmas manuscritas para, posteriormente, entrenar clasificadores para cada usuario.

Debido a que en los sistemas reales la lista de usuarios no es fija, se considera un conjunto desunido de usuarios, uno para aprender las características y otro para entrenar los clasificadores WD. De este modo se verifica que el extractor de características generaliza para nuevos usuarios.

Se entrena la CNN con el **conjunto de desarrollo** de firmas D utilizando las formulaciones definidas en la Tabla 2.3. Esto permite crear un modelo extractor de características. Posteriormente, se utiliza el **conjunto de explotación** E para entrenar un clasificador binario para cada usuario SVM (Support Vector Machines), introduciendo como entrada las características extraidas gracias al modelo creado por la CNN. La hipótesis es que las firmas genuinas y las falsificaciones serán más fáciles de separar en este espacio de características, donde, si la red tiene éxito, se habrán capturado las propiedades intrínsecas de las firmas. Por lo tanto, las redes neuronales convolucionales son una arquitectura particularmente adecuada para la verificación de firmas.

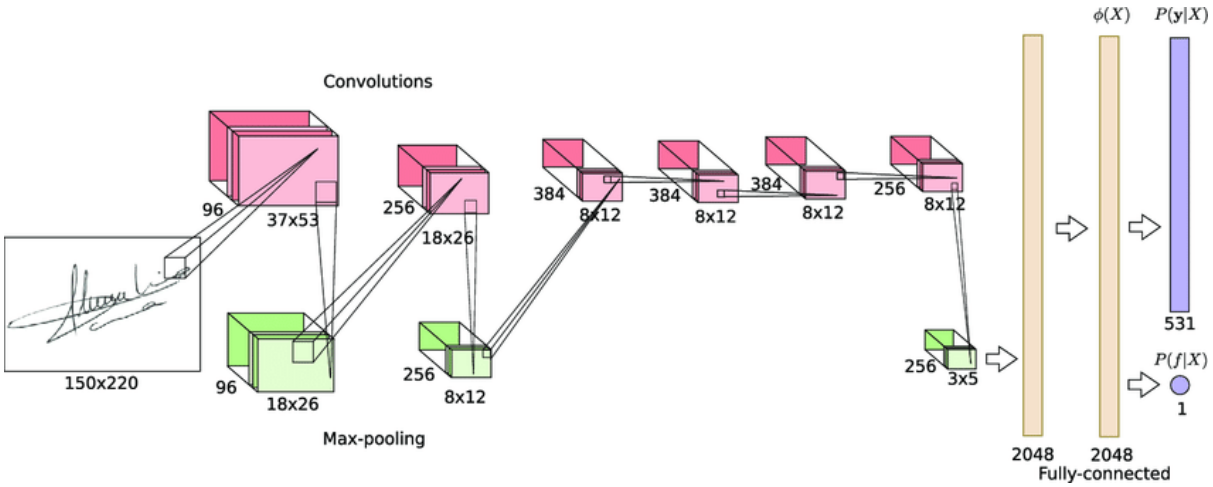

Figura 2.5: Ilustración de la estructura de la CNN usada. Fuente: [1]

Este tipo de arquitectura, ver Fig.2.5, se adapta mejor que los modelos totalmente conectados para tamaños de entrada grandes, tener un menor número de parámetros de entrada es una propiedad deseable para el problema en cuestión, ya que no podemos reducir las imágenes de las firmas demasiado sin el riesgo de perder los detalles que permiten discriminar entre falsificaciones especializadas y firmas genuinas (por ejemplo, la calidad de los trazos de la pluma).

La idea clave detrás de este enfoque es enseñar a la red a distinguir entre usuarios, esperando que aprenda una jerarquía de representaciones, y que las representaciones en las últimas capas capture las propiedades relevantes de las firmas. En particular, si la red logra distinguir entre los diferentes usuarios del conjunto de desarrollo, luego la representación de las firmas de estos usuarios serán linealmente separables en el espacio de representación, ya que la última capa es un clasificador lineal con respecto a su entrada.

En el el sistema de verificación propuesto por la Universidad de Québec y la de Paraná se utiliza para el desarrollo del sistema la base de datos GPDS [24] siguiendo la división descrita en la Fig. 2.6

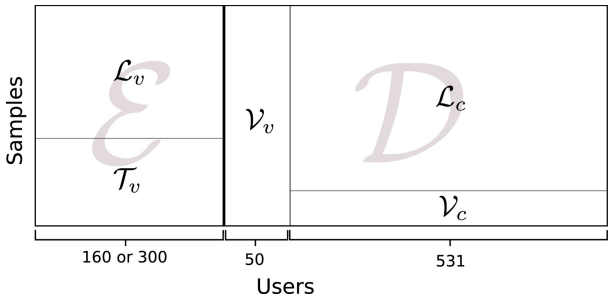

Figura 2.6: La base de datos GPDS es separada en un conjunto de explotación y un conjunto de desarrollo. El conjunto de desarrollo se usa para el aprendizaje de las características y hacer todas las decisiones del modelo. El conjunto de explotación representa los nuevos usuarios que llegan al sistema, en donde se entrenan los clasificadores WD solo usando firmas genuinas. Fuente: [1]

El sistema se divide en las siguientes etapas: preprocesado, entrenamiento de la red neuronal convolucional y, por último, entrenamiento de los clasificadores WD.

Por último, una vez creados los modelos con el conjunto de desarrollo de la base de datos GPDS, utilizan conjuntos de explotación de otras bases de datos para ver como generaliza el sistema.

Preprocesado

Las firmas de los conjuntos de datos utilizados para este sistema [1] se extraen de los documentos donde se escribieron, por lo que la extracción de la firma no se trata.

Sin embargo, se requieren pasos de preprocesamiento. Las redes neuronales esperan entradas de un tamaño fijo, y dado que en algunas bases de datos las firmas varían significativamente es necesario la fase de preprocesamiento. Primero se centran las firmas en un gran lienzo de tamaño Scanvas = H x W, usando el centro de masa de las imágenes. Se elimina el fondo utilizando el algoritmo de OTSU, configurando el fondo a blanco (intensidad 255), y dejando los píxeles del primer plano en escala de grises. La imagen es entonces invertida al restar cada píxel el brillo máximo I (x, y) = 255 - I (x, y), dando valor cera al fondo. Por último, la imagen es redimensionada al tamaño de entrada de la red.

Entrenamiento de la Red Neuronal Convolucional

Esta sección describe los detalles de las diferentes capas de la CNN que se pueden ver en detalle en la Tabla 2.3.

Cuenta con una primera capa, **la capa de entrada**, donde el tamaño de las imágenes es de 1 x 150 x 220 al ser imágenes en escala de grises. Después hay cinco **capas convolucionales** que

Layer	Size	Other Parameters
Input	1x150x220	
Convolution (C1)	96x11x11	stride = 4, pad=0
Pooling	96x3x3	stride = 2
Convolution (C2)	256x5x5	stride = 1, pad=2
Pooling	256x3x3	stride = 2
Convolution (C3)	384x3x3	stride = 1, pad=1
Convolution (C4)	384x3x3	stride = 1, pad=1
Convolution (C5)	256x3x3	stride = 1, pad=1
Pooling	256x3x3	stride = 2
Fully Connected (FC6)	2048	
Fully Connected (FC7)	2048	
Fully Connected + Softmax ($P(\mathbf{y}\|X)$)	M	
Fully Connected + Sigmoid ($P(f\|X)$)	1	

Cuadro 2.3: Resumen de las capas de la CNN. Fuente: [1]

se intercalan con **capas de pooling** en donde el tamaño de los datos es reducido hasta llegar a **las capas Fully Connected** de tamaño 2048. Es decir, se consigue reducir las características de la imagen en un vector de 1 x 2048. Estas últimas características son la que se utilizarán en el SVM específico para cada usuario al introducir una imagen de una firma de un nuevo usuario, una vez entrenada la red. La red neuronal cuenta al final con **la capa Softmax** para realizar el entrenamiento en donde, a través de las etiquetas, va distribuyendo los pesos para saber si el entrenamiento se realiza correctamente. Es decir, para saber a qué usuario pertenece la imagen introducida, ya que M es el número de los usuario utilizados en el conjunto de desarrollo para entrenar la red y devuelve la probabilidad asignada por la red a cada posible usuario, siendo correcta si ésta es la más alta para el usuario en cuestión. La última capa de todas, **la capa Sigmoid**, permite a la red saber si la firma es genuina o falsificación, dando una probabilidad de que la firma pertenezca o no al usuario.

En la Tabla 2.4 se pueden ver las operaciones implementadas de prealimentación de cada capa.

Table 2
List of feedforward operations.

Operation	Formula
Convolution	$\mathbf{z}^l = \mathbf{h}^{l-1} * W^l$
MaxPooling	$h^l_{xy} = \max_{i=0,\dots,s, j=0,\dots,s} \mathbf{h}^{l-1}_{(x+i)(y+j)}$
Fully-connected layer	$\mathbf{z}^l = W^l \mathbf{h}^{l-1}$
ReLU	$\mathrm{ReLU}(z_i) = \max(0, z_i)$
Sigmoid	$\sigma(z_i) = \frac{1}{1+e^{-z_i}}$
Softmax	$\mathrm{softmax}(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$
Batch Normalization	$\mathrm{BN}(z_i) = \gamma_i \hat{z}_i + \beta_i,$ $\hat{z}_i = \frac{z_i - \mathrm{E}[z_i]}{\sqrt{\mathrm{Var}[z_i]}}$

$\mathbf{z}^l$: pre-activation output of layer l
$\mathbf{h}^l$: activation of layer l
$*$: discrete convolution operator
W, γ, β: learnable parameters

Cuadro 2.4: Lista de las operaciones de prealimentación. Fuente: [1]

Una vez diseñada la red neuronal, existen dos modos de entrenarla: usando solamente las firmas genuinas de los usuarios o usando también las falsificaciones de estos. Por lo tanto, podemos llegar a dos modelos diferentes de los cuales extraer características de los usuarios del conjunto de explotación.

- **Modelo SigNet**: se extrae entrenando solamente la red neuronal con firmas genuinas de los usuarios del conjunto de desarrollo.
- **Modelo SigNet-F**: se extrae entrenando la red neuronal con las firmas genuinas y falsificaciones de los usuarios del conjunto de desarrollo.

Entrenamiento de los Clasificadores Dependientes del Usuario

Después de entrenar la CNN y extraer las 2048 características para cada firma de cada usuario. Se utiliza un SVM específico para cada usuario en el conjunto de explotación, el cuál no ha sido usado en ningún momento para entrenar la red convolucional

Para cada usuario se construye un conjunto de capacitación que consiste en firmas genuinas del usuario como **muestras positivas** y firmas genuinas de otros usuarios como **muestras negativas**. De este modo, entrenamos las Máquinas de Soporte de Vectores (SVM) usando una ponderación diferente para la clase positiva y negativa para explicar el desequilibrio de tener muchas más muestras negativas que positivas. Se puede utilizar tanto una formulación lineal (Linear SVM) como función de base radial (RBF Kernel). Para las pruebas de test se utiliza un conjunto separado de firmas genuinas del usuario (es decir, no se utilizan en el entrenamiento del SVM) y un conjunto de muestras negativas que pueden ser de firmas de otros usuarios, denominadas **random forgeries** (no utilizadas en el entrenamiento del SVM) o falsificaciones, denominadas **skilled forgeries**.

Protocolo Experimental

A continuación, se describe el protocolo experimental utilizado para la base de datos MCYT-75 [25] ya que será la utilizada para comparar los resultados con el sistema propuesto en este trabajo.

MCYT-75 es un subconjunto de usuarios de la base de datos MCYT, la cuál cuenta tanto con las firmas off-line como las on-line. Es por esto que podemos llegar a comparar los resultados obtenido, ya que para este trabajo se hace uso de la firma off-line sintética obtenida de las firmas on-line.

MCYT-75 cuenta con 15 firmas **genuinas** y 15 firmas **skilled** (falsificaciones) por usuario.

Los 75 usuarios son el conjunto de explotación, en donde una vez extraídas las características a través de la red neuronal de los modelos descritos anteriormente (SigNet y SigNet-F), se entrenan los SVM. Se utilizan 10 firmas genuinas del usuario como muestras positivas y 10 firmas del resto de usuarios como muestras negativas para el conjunto de entrenamiento del SVM por usuario. Para el conjunto de test se utilizan las otras 5 firmas genuinas del usuario no utilizadas en el entrenamiento y las 15 firmas falsificadas (skilled forgeries) de ese usuario. De este modo, no se tienen en cuenta la identificaión del usuario, solamente si la firma es auténtica o falsificación. Por lo tanto, el resultado expresa lo bueno que es el sistema para **skilled forgeries**.

Resultados

Cabe destacar que los resultados obtenidos son de la firma estática original, y en este trabajo se trabajará con firma estática sintética. Los resultados están expresados en **Equal Error Rate** visto anteriormente, y sobre las **skilled forgeries**, es decir, sobre las falsificaciones de los usuarios.

Reference	**# Samples**	**Features**	**EER**
Gilperez et al. [26]	5	Contours (chi squared distance)	10.18
Gilperez et al. [26]	10	Contours (chi squared distance)	6.44
Wen et al. [27]	5	RPF (HMM)	15.02
Vargas et al. [28]	5	LBP (SVM)	11.9
Vargas et al. [28]	10	LBP (SVM)	7.08
Ooi et al [29]	5	DRT + PCA (PNN)	13.86
Ooi et al [29]	10	DRT + PCA (PNN)	9.87
Soleimani et al.[30]	5	HOG (DMML)	13.44
Soleimani et al. [30]	10	HOG (DMML)	9.86
Proposed	5	SigNet (SVM)	3.58 ($\pm$ 0.54)
Proposed	10	SigNet (SVM)	**2.87 ($\pm$ 0.42)**
Proposed	5	SigNet-F (SVM)	3.70 ($\pm$ 0.79)
Proposed	10	SigNet-F (SVM)	3.00 ($\pm$ 0.56)

Cuadro 2.5: Comparación con el estado del arte en MCYT. EER de skilled forgeries. Fuente: [1]

Como se puede apreciar en la Tabla 2.5, el sistema mejora con creces el estado del arte, llegando a un resultado de **2.87 ($\pm$ 0.42)** utilizando 10 firmas por usuario y el modelo **SigNet**. Éste no utiliza en ningún momento las falsificaciones de los usuarios para el entrenamiento y de todos modos obtiene unos buenos resultados a la hora de distinguir entre firmas genuinas y falsificaciones.

3

Sistema propuesto

3.1. Introducción

A continuación, se va a desarrollar paso a paso todo el proceso seguido para la creación del sistema propuesto. Desde el entrenamiento de la CNN hasta el entrenamiento de los clasificadores SVM Writer-Dependent (WD).

Cabe destacar que todo el sistema es entrenado utilizando solamente firmas genuinas, tanto la CNN como el SVM. Es decir, el sistema es capar de diferenciar entre una firma genuina y una falsificación (skilled forgery) sin haber tenido contacto con estas últimas en ningún momento. Esto es muy importante para las aplicaciones reales del sistema, ya que si llega una nueva base de datos, no tiene porqué contar con firmas falsificadas para cada usuario.

A su vez, el sistema utiliza firmas off-line sintéticas, extraídas de las firmas on-line como se explicará en la Sección 4.

En la Sección 5. Se pondrá a prueba la robustez del sistema utilizando tanto un número pequeño de usuarios, como pocas firmas genuinas de éstos.

3.2. Entrenamiento de la CNN

La finalidad de entrenar una Red Neuronal Convolucional es crear un modelo que extraiga las características de los nuevos usuarios para luego utilizarlas en el entrenamiento de los clasificadores.

Cuanto más robusto sea el modelo, mejor preparadas estarán las características para verificar e identificar nuevos usuarios, así como para poder generalizar a otras bases de datos y entornos. El estudio de la decisión de estructura de la base de datos viene dado por la Universidad de Québec y la de Paraná, como se ha expuesto en la Sección 2.4.4. [1].

La principal diferencia es la necesidad de crear un modelo específico para firmas off-line sintéticas, extrayendo las imágenes de los datos de la firma on-line.

La arquitectura en detalle se puede ver en la Tabla. 2.3. Y fue la llevada en este sistema también debido a sus buenos resultados. El trabajo de mejora de la red neuronal se deja para futuros estudios, ya que en éste se centra más en la creación inicial del sistema.

La optimización de la red se llevó a cabo minimizando la pérdida con la función de descenso de gradiente estocástico con Nesterov Momentum [1], un factor de impulso de 0.9 y utilizando minibatches de tamaño 32. Los modelos fueron entrenados para 60 épocas (ajustes de los pesos para todos los pares de entrenamiento), con una tasa de aprendizaje inicial de 10^{-3}, que se dividió por 10 cada 20 épocas. También se usa el aumento de datos (data augmentation) para aumentar el número de imágenes y evitar el sobre ajuste de la red (overfitting), mediante el uso de recortes aleatorios de tamaño 150 × 220 de la imagen original de 170 × 242. Por último, se aplica una normalización de los batches (batch normalization), calculada a partir de cada uno de los minibatches para facilitar el entrenamiento de la red.

Los datos son divididos en dos sets:

- **Set de Entrenamiento**: el 90 % de los usuarios. Utilizado para crear el modelo.
- **Set de Test**: el 10 % de los usuarios. Utilizado para validar el modelo.

En cada época se crea un nuevo modelo con los datos de entrenamiento que se valida y escoge los parámetros que mejor se ajustan a los datos de test, ya que de ambos la red tiene las etiquetas y sabe la exactitud alcanzada. De este modo, la red es capaz de extraer el modelo que mejor se ha ajustado a los datos de entrenamiento y test a lo largo de las 60 épocas.

3.3. Entrenamiento de los clasificadores Writer-Dependent

A continuación, se explicará en detalle el funcionamiento del clasificador SVM (Support Vector Machine)

Las **Máquinas de vectores de soporte** son un método de aprendizaje automático para resolver problemas de clasificación y regresión. El algoritmo crea un hiperplano que separa los datos en clases gracias a vectores de soporte. Los vectores de soporte son los puntos más cercanos al hiperplano que permiten calcular la distancia a éste. De este modo, maximizando esta distancia se podrá clasificar de manera más efectiva los datos.

En este sistema el aprendizaje es supervisado, ya que se cuenta con las etiquetas de las firmas que entran al sistema, tanto genuinas como falsificaciones.

La finalidad es crear un SVM por usuario (writer-dependent), en donde a través de muestras positivas (firmas genuinas del usuario) y muestras negativas (firmas genuinas de otros usuario) el clasificador pueda separar de la manera más óptima ambas clases, para que cuando lleguen firmas nuevas del usuario o falsificaciones se verifique la identidad de éste.

Se introducen las 2048 características extraídas por el modelo de la CNN para cada firma, y es gracias a estos datos que el SVM clasifica a los usuarios. A continuación se muestran las diferentes formas de entrenar y testear el clasificador:

1. **Fase de Entrenamiento**: en la fase de entrenamiento se introducen como muestras positivas las firmas genuinas del usuario. Este número puede varias dependiendo de la base de datos y la finalidad del sistema. Como muestras negativas se pueden introducir firmas genuinas de usuario del conjunto de desarrollo (usador para entrenar la CNN) y/o firmas genuinas de otros usuarios del conjunto de explotación, ambas son denominadas **random forgeries**. A su vez, se pueden entrenar también con falsificaciones del usuario, **skilled forgeries**, o de otros usuarios.

2. **Fase de Test**: en la fase de test se reconoce la precisión del sistema. Se introducen firmas genuinas del usuario no usadas en la fase de entrenamiento, random forgeries y skilled forgeries. De este modo podemos comprobar cuánto de bueno es el sistema tanto como para identificar el usuario entre la base de datos como para verificar que no se trata de una falsificación.

El SVM devuelve una puntuación para cada firma, indicando la distancia hasta el hiperplano. De este modo se puede determinar en qué lado de éste se encuentra y si corresponde con una firma genuina del usuario o no. Una vez entrenado el clasificador se explican los diferentes **Equal Error Rates** que se utilizarán en la parte experimental del trabajo:

- **EER Random (global threshold)**: EER calculado con las firmas genuinas y las random forgeries introducidas en la fase de test. Se extraen las puntuaciones de todos los usuarios del conjunto de explotación y se calcula el error gracias a un umbral global.
- **EER Random (user thresholds)**: EER calculado con las firmas genuinas y las random forgeries introducidas en la fase de test. Se extraen las puntuaciones para cada usuario y se calcula el error por usuario gracias a un umbral específico para cada uno, después se calcula la media entre todos.
- **EER Skilled (global threshold)**: EER calculado con las firmas genuinas y las skilled forgeries introducidas en la fase de test. Se extraen las puntuaciones de todos los usuarios del conjunto de explotación y se calcula el error gracias a un umbral global.
- **EER Skilled (user thresholds)**: EER calculado con las firmas genuinas y las skilled forgeries introducidas en la fase de test. Se extraen las puntuaciones para cada usuario y se calcula el error por usuario gracias a un umbral específico para cada uno, después se calcula la media entre todos.

4

Bases de Datos

4.1. Introducción

Hoy en día se disponen de varias bases de datos utilizadas en diferentes estudios del estado del arte, lo cual permite la comparación de los algoritmos y sistemas implementados. La carencia de bases de datos públicas debida a los problemas legales, a la privacidad de los usuarios y a las limitaciones tecnológicas para conseguir firmas tanto on-line como off-line de los últimos años ha sido subsanada debido al desarrollo tecnológico que ha permitido conseguir una gran cantidad de información de firma manuscrita, suficiente como para conseguir resultados competitivos comparados con otros rangos biométricos más extendidos en la actualidad.

Las bases de datos son la clave de los sistemas basados en redes neuronales ya que se necesitan datos reales para poder desarrollar sistemas de buen nivel de verificación de firma.

El proceso de obtención de **DeepSignDB (DeepSign Database)** viene detallado explicado en el trabajo fin de grado de Pablo Lázaro [2]. Ya que este trabajo se centra en la creación de un sistema de verificación de firma off-line sintética a través de esta base de datos.

Esta base de datos será de gran importancia para el segundo experimento de este trabajo que trata de crear un modelo global. Expuesto en la Sección 5.3.

4.2. DeepSignDB

La base de datos DeepSignDB es la suma de diferentes bases de datos, creando así una más extensa. Se han utilizado las siguientes bases de datos públicas para su creación: MCYT, BiosecurID, Biosecure DS2, e-BioSign DS1, e-BioSign DS2 y e-BioSign DS3. A continuación, se describirán las características principales de cada base de datos: año de captura, número de usuarios, número de sesiones, número de firmas genuinas y falsificadas (*skilled forgeries*) y dispositivos y útiles de escritura utilizados, ver Tabla 4.1.

- **MCYT** [25]: la adquisición fue llevada a cabo por diversas instituciones universitarias españolas en el año 2003, entre las que se encuentra el Grupo de Reconocimiento Biométrico ATVS, de la Universidad Autónoma de Madrid. Esta base de datos cuenta con un total

	MCYT	BiosecurID	BiosecureDS2	e-BioSign DS1	e-BioSign DS2-DS3
Año	2003	2007	2008	2016	2016-2017
Usuarios	330	400	676	65	81
Sesiones	1	4	2	2	2
#muestras genuinas/ usuario/dispositivo	25	16	30	8	8
#falsificaciones/ usuario/dispositivo	25	12	20	6	6
Dispositivo (Útil de escritura)	W. Intous (st)	W. Intuos3 (st)	W. Intuos3 (st)	W. STU-50 (st) W. STU-53 (st) W. DTU-1031 (st) SG. Gal.Note(st/d) SG. ATIV7 (st/d)	W. STU-530 (st) SG. Gal.Note (st/d) SG. Galaxy Neo S3 (d)

Cuadro 4.1: Características de las bases de datos utilizadas. SG=Dispositivo Samsung, W=Tablet Wacom, st=stylus, d=dedo. Fuente: [2]

de 330 usuarios. Las firmas fueron capturadas usando un dispositivo WACOM Intuos A6 tablet con una frecuencia de muestreo de 100 Hz, permitiendo capturar las coordenadas, la presión y los ángulos del bolígrafo (azimuth y altitud). Hay 25 firmas genuinas y 25 firmas falsificadas por usuario. Las firmas fueron capturadas en grupos de 5. Primero 5 firmas genuinas, luego 5 firmas falsificadas de otro usuario, repitiendo este procedimiento hasta alcanzar las 25 firmas de cada tipo. Cada usuario proporciona 5 firmas falsificadas para los 5 usuarios previos en la base de datos.

- **BiosecurID** [31]: trata de un proyecto financiado por el Ministerio de Ciencia y Tecnología en el cual han participado seis instituciones académicas españolas entre las que se encuentra el Grupo de Reconocimiento Biométrico ATVS de la UAM. número de usuarios es de 400 y consta de 4 sesiones llevadas a cabo en intervalos de un mes. Una característica de esta base de datos es la distribución balanceada en la edad de los usuarios que participaron, contando con usuarios entre 18 y más de 45 años. Las firmas fueron capturadas con una WACOM Intuos3 A4. En cada una de las 4 sesiones, cada usuario realiza 4 firmas genuinas y 1 firma falsificada para cada uno de los 3 usuarios anteriores. Se consideran 4 escenarios de falsificación.

- **Biosecure DS2** [32]: en este proyecto participan más de 30 instituciones de investigación procedentes de 15 países diferentes. El Grupo de Reconocimiento Biométrico ATVS también participó en la elaboración de dicha base de datos. Fue capturada utilizando una WACOM Intuos3 A6 digitalizada a una frecuencia de muestreo de 100 Hz, con un procedimiento similar al seguido en la base de datos MCYT. El dispositivo captura información de coordenadas de posición, presión y ángulos de inclinación del bolígrafo (azimuth y altitud). Cuenta con un total de 667 usuarios, de 7 países diferentes. Por cada uno de los usuarios se posee un total de 30 firmas genuinas y 20 falsificadas, capturadas en 2 sesiones con un espacio temporal de unos 2 meses. Las firmas fueron capturadas en bloques de 5. En cada sesión los usuarios realizaron 3 sets de 5 firmas genuinas y 5 firmas falsificadas entre cada set. Cada usuario realizó 5 falsificaciones para los 4 usuarios anteriores de la base de datos. El usuario tenía acceso visual a la información dinámica de la firma a falsificar.

- **e-BioSign DS1** [21]: la idea de crear esta base de datos surge al intentar abordar los problemas que supone entrenar y evaluar un sistema con múltiples dispositivos y útiles de

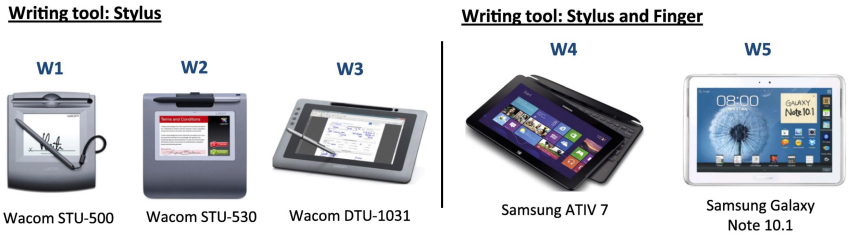

Figura 4.1: Dispositivos de captura utilizados en e-BioSign DS1. Fuente: [18].

escritura. La adquisición fue llevada a cabo por el grupo de Reconocimiento Biométrico ATVS exclusivamente. Esta base de datos está compuesta por 65 usuarios, cuyos datos se recopilan en dos sesiones, con un periodo de tiempo entre ellas de 3 semanas. La base de datos está compuesta por cinco dispositivos de captura de escritura a mano. Tres de ellos están diseñados específicamente para esta tarea (dispositivos Wacom), mientras que los otros dos son tabletas de uso general (tabletas Samsung). Vale la pena señalar que los cinco dispositivos se usaron con su propio lápiz óptico y que en los dos dispositivos Samsung se usó también el dedo como herramienta de escritura, ver Tabla 4.1, lo que va a permite analizar el efecto de la herramienta de escritura en el rendimiento del sistema. Además, se utilizó el mismo protocolo de captura para los cinco dispositivos. Los dispositivos Wacom utilizados son los siguientes: Wacom STU-500 (**W1**), Wacom STU-530 (**W2**) y Wacom DTU-1031 (**W3**). Cabe destacar que la frecuencia de muestreo es la misma para todas las Wacom (200Hz) y los niveles máximos de presión varían entre 1024 para STU-530 y 512 para STU-500 y DTU-1031. Los dispositivos genéricos utilizados son: Samsung ATIV 7 (**W4**) y Samsung Galaxy Note 10.1 (**W5**), ver Fig. 4.1. Ambos tienen 1024 niveles de presión. Para cada usuario se capturaron en cada dispositivo 8 firmas genuinas y 6 falsificadas, es decir, 3 genuinas y 2 falsificadas en cada una de las dos sesiones.

- **e-BioSign DS2-DS3**: esta base de datos fue capturada entre los años 2016 y 2017 por el grupo de Reconocimiento Biométrico ATVS en la Universidad Autónoma de Madrid. En primer lugar, se creó solamente la base DS2 con 53 usuarios en 2016 y en 2017 se introdujeron 28 usuarios más, formando un total de 91 usuarios. Los tres dispositivos que han intervenido en esta base de datos han sido una tableta WACOM-STU530 (**W2**), una tableta Samsung Galaxy Note 10.1 (**W5**) y un Smartphone Samsung Galaxy Neo SIII (**W6**). Las firmas fueron realizadas con stylus para W2, con stylus y dedo para W5 y con dedo para W6. En cada dispositivo se realizaron dos sesiones con un intervalo de 15 días entre ellas. El procedimiento para la obtención de genuinas y falsificadas es similar al de e-BioSign DS1.

4.3. Preprocesado de los datos de firma on-line

Antes de extraer la firma off-line es necesario realizar un preprocesado de las firmas on-line para evitar firmas vacías, errores de captura, etc.

Como esto ya ha sido realizado no se va a entrar en detalle. En resumidas cuentas, se hizo el siguiente proceso para todas bases de datos:

1. Recorrer todos los usuarios para comprobar que tienen el número de firmas genuinas y falsificadas que corresponde.

2. Búsqueda de firmas vacías en todos los usuarios, es decir, firmas cuyos ficheros no tienen ningún valor almacenado.

3. Conversión de las firmas del formato de origen (i.e. .svc, .fpg) a un formato común para su uso posterior (i.e. .txt).

Tras esto se eliminaron 36 usuarios de la base de datos Biosecure DS2 ya que no cumplían con los requisitos expuestos anteriormente.

Como se muestra en la Fig.4.2, se obtiene la firma off-line a partir de la on-line, utilizando los datos de las coordenadas y la presión. Las imágenes se obtienen mediante Matlab y con formato .jpg.

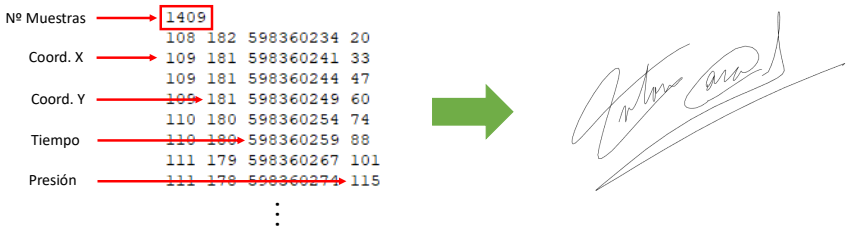

Figura 4.2: Elementos de la firma on-line (izquierda) y su correspondiente firma off-line(derecha). Fuente: [2]

Dentro de los parámetros que se pueden variar a la hora de extraer la firma sintética se destacan los siguientes:

- **Tamaño del píxel**: se puede aumentar o disminuir el tamaño del píxel para que se asemeje lo más posible a la firma original.
- **Información de vuelo**: se puede dibujar la información de vuelo (pen-ups) del usuario ya que los dispositivos de capturas guardan estos valores.
- **Presión**: se puede hacer un escalado de la presión de cada muestra y asignarle un valor de la escala de grises, de este modo asemejándose a la presión que haría un usuario al apretar con la pluma en el dispositivo o en el papel. Este parámetro es de gran importancia ya que varía mucho de un usuario a otro y nos permite diferenciar con mayor facilidad a éstos.

4.4. Preprocesado de las imágenes de firma off-line

Una vez que hemos extraído las imágenes de las firmas para cada usuario es necesario (como se comentó en la Sección 2.4.4.) realizar un preprocesado de las imágenes antes de introducirlas en la red neuronal.

Dado que las imágenes al ser sintéticas ya están en escala de grises y centradas por su centro de masas no es necesario aplicar el algoritmo OTSU ni el centrado. Simplemente es necesario hallar la imagen invertida y redimensionarla al tamaño de entrada de la red.

Para facilitar el proceso posterior de entrenar la CNN y los clasificadores se crea un fichero *numpy .npz* en Python que engloba todas las firmas de todos los usuarios de la base de datos. El cuál contiene para cada usuario:

- **x**: np.ndarray (N x 1 x H x W) donde N son el número de imágenes del usuario, en escala de grises de tamaño H x W
- **y**: np.ndarray (N) indicando el nombre del usuario que escribió esa firma, la etiqueta.
- **yforg**: np.ndarray (N) indicando si la firma es genuina o falsificación.
- **usermapping**: utilizado para el mapeado de índices en Python
- **filenames**: np.ndarray(str) (N), el nombre de fichero asociado al usuario y a sus firmas

Cabe destacar que por motivos de reserva de memoria, es necesario que para una misma base de datos todos los usuarios tengan el mismo número de firmas tanto genuinas como skilled.

Por lo tanto, al ser **DeepSignDB** una base de datos compuesta por otras, es conveniente globalizarla para que todos los usuarios tengan el mismo número de firmas.

4.5. Base de Datos definitiva

Debido a que para el correcto funcionamiento del sistema es necesario que la base de datos cumpla con que todos los usuarios tengan el mismo número de firmas, tanto genuinas como falsificación (explicado en la Sección anterior) es conveniente hacer una toma de decisión.

Al ser **BioSecurID** la base de datos con un número menor de firmas genuinas por usuario, ver Tabla 4.1, se tomará éste como referencia para **MCYT** y **BioSecure DS2**. Tanto **e-BioSign DS1** como **e-BioSign DS2-DS3** no se usarán para la base de datos definitiva ya que son de dedo, y a su vez tienen un número de usuarios menor y no cumplen con los requisitos del número de firmas genuinas por usuario. De todos modos, estas bases de datos se pueden utilizar para evaluar la generalización del sistema en futuros trabajos.

Por lo tanto, se cogerán 16 firmas genuinas por usuario y 12 falsificaciones, ver Tabla 4.2.

Nombre de la Base de Datos	**# genuinas**	**# skilled**	**Total**
MCYT	16	12	28
BioSecure DS2	16	12	28
BioSecurID	16	12	28

Cuadro 4.2: Distribución de la Base de Datos en número de firmas por usuario. Genuinas y falsificaciones.

El proceso de selección fue intentando imitar las sesiones de BioSecurID, que se divide en 4 sesiones de 4 firmas cada una para las firmas genuinas. Y se usan las 3 primeras para el entrenamiento y la última para la evaluación. Es por esto, que se eligen las 12 primeras firmas de cada usuario de MCYT y BioSecure DS2 para entrenamiento y las 4 últimas para evaluación.

4.6. División de la Base de Datos

Es necesario hacer una división de los usuarios en el conjunto de desarrollo, necesario para entrenar la CNN y el conjunto de explotación, utilizado para entrenar los clasificadores WD y evaluar el sistema.

Como la principal finalidad del sistema es su robustez y capacidad para generalizar a otras bases de datos es conveniente que ambos conjuntos contengan firmas de las tres bases de datos. En la Tabla 4.3 se puede ver la división que se ha llevado a cabo.

Nombre de la BBDD	Desarrollo	Explotación	Total
MCYT	230	100	330
BioSecure DS2	510	140	650
BioSecurID	268	132	400
Total	**1008**	**372**	**1380**

Cuadro 4.3: Distribución de la Base de Datos en el Conjunto de Desarrollo y Explotación

- **MCYT**: en total se han cogido 330 usuarios, 230 para desarrollo y 100 para explotación. La razón de elegir 100 usuarios para explotación es porque 145 usuarios fueron capturados por el grupo de Reconocimiento Biométrico ATVS, por lo tanto se pueden publicar los resultados al ser una base de datos totalmente pública, eligiendo 100 de ellos.
- **BioSecure DS2**: esta base de datos contiene un total de 650 usuarios, por lo tanto, será la que más aporte a nivel de número de usuarios en la base de datos final. De los 145 usuarios que capturó ATVS, se eligieron 140 usuarios de explotación. Y el resto, 510 usuarios para desarrollo.
- **BioSecurID**: de los 400 usuarios que contiene esta base de datos, 286 han sido seleccionados para desarrollo y 132 para explotación, ya que estos 132 usuarios han sido capturados por el grupo de Reconocimiento Biométrico ATVS.

Por lo tanto, el sistema utiliza 1008 usuarios en la etapa de desarrollo y 372 en la etapa de explotación.

5

Desarrollo experimental

En este capítulo se exponen los diferentes experimentos realizados en el sistema propuesto de redes neuronales convolucionales. Para cada experimento se explica en detalle el protocolo experimental seguido que permita comparar con el estado del arte o con futuros experimentos. Este trabajo se centra en la obtención de la mejor versión posible de firma sintética para su posterior uso en la creación de un modelo global. Esto se debido a que el modelo existente estaba diseñado para firmas estáticas originales y no sintéticas. Por último, se tratan los resultados para distintas bases de datos y se hará una comparación con el estado del arte.

5.1. Creación de la arquitectura y adaptación al entorno

Para poder desarrollar el sistema y los experimentos fue necesario crear un entorno de programación específico que cumpliese con todos los requerimientos necesarios para la programación de aprendizaje automatizado.

Para la extracción de las firmas off-line sintéticas se utiliza MatLab, a través de códigos utilizados en trabajos anteriores [2], ajustándolos a las bases de datos y al sistema.

El sistema propuesto está programado en Python (versión 3.6) a través de Anaconda. Anaconda es una herramienta que te permite crear entornos específicos orientados a simplificar el despliegue y la administración de los paquetes de software. Se utilizan funciones originales de PyTorch para el entrenamiento de la CNN y funciones de Scikit-learn para el entrenamiento de los SVM.

5.2. Obtención de la mejor versión de firma off-line sintética

En este experimento se realizan diferentes versiones de la firma sintética para decidir cuál es la que mejor funciona y se adapta al sistema.

5.2.1. Protocolo experimental

A continuación, se expone el protocolo experimental que nos permite comparar con el estado del arte para la base de datos MCYT-75 [25].

MCYT-75 cuenta con 72 usuarios de los 75 que tienen tanto la versión off-line original como la versión on-line. De este modo nos permite extraer las firmas sintéticas y comparar con las originales. De ahora en adelante la denominaremos MCYT-72. Ésta consta con 15 firmas genuinas y 15 falsificaciones para cada usuario.

Para la creación de los modelos específicos se utilizará la base de datos completa de MCYT que cuenta con 330 usuarios. Se extraen los 72 usuarios que aparecen en MCYT-72 y son utilizados para los clasificadores, siendo el conjunto de explotación. De este modo quedan 258 usuarios para el entrenamiento de la CNN, conjunto de desarrollo. En la Tabla 5.1 se puede ver esta división.

Nombre de la BBDD	**Desarrollo**	**Explotación**	**Total**
MCYT	258	72	330

Cuadro 5.1: División de la base de datos MCYT.

Se utilizará tanto el modelo **SigNet** [1] como **modelos específicos** para la extracción de características. El entrenamiento de los SVM se puede ver en la Tabla 5.2.

	Entrenamiento		**Test**		
	Genuinas	**Random**	**Genuinas**	**Random**	**Skilled**
MCYT-72	10	10 x 71 = 710	5	5	15

Cuadro 5.2: División del SVM para cada usuario de MCYT-72.

5.2.2. Desarrollo experimental

Creación de las diferentes versiones de MCYT

Se crean 6 versiones de la base de datos MCYT completa de las cuales se extraerá un modelo diferente para cada una con el que extraer las características del conjunto de explotación:

- **Versión 1**: Se utiliza un tamaño de píxel de 1.
- **Versión 2**: Se utiliza un tamaño de píxel de 2.
- **Versión 3**: Se utiliza un tamaño de píxel de 3.
- **Versión 4**: Se utiliza un tamaño de píxel de 4.
- **Versión 5**: Se utiliza un tamaño de píxel de 3 y además se muestra la información de vuelo (pen-ups).
- **Versión 6**: Se utiliza un tamaño de pixel de 3 y además se muestra la información de presión.

El proceso de extracción de la firma sintética viene explicado en detalle en la Sección 4.3. y 4.4.

Cabe destacar que la versión 5 y 6 fueron creadas con un tamaño de pixel 3 ya que es la que daba mejores resultados como se verá más adelante. Es decir, primero se decidió el tamaño de pixel y después se planteó añadir la información de vuelo y presión.

En la Figura 5.1 se muestra una comparación de la variación del tamaño del pixel con una firma original. De modo subjetivo la versión 3 (tamaño de pixel 3) es la que más se asemeja a la firma original, de todos modos se comprobará en los resultados objetivamente.

Y en la Figura 5.2 se muestra una comparación de añadir la información de vuelo, y de añadir la de presión con una firma original. Esta última aumenta aún más la semejanza a la firma original ya que imita la presión causada por el usuario en el dispositivo de reconocimiento.

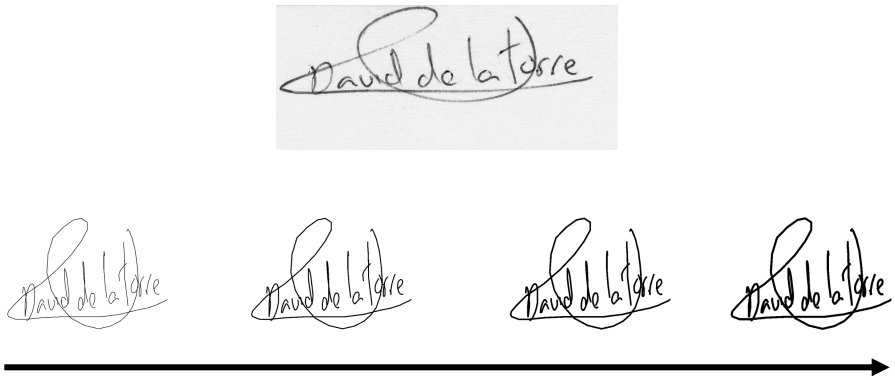

Figura 5.1: Comparación de la firma sintética con la firma original. De izquierda a derecha: versión 1, versión 2, versión 3 y versión 4.

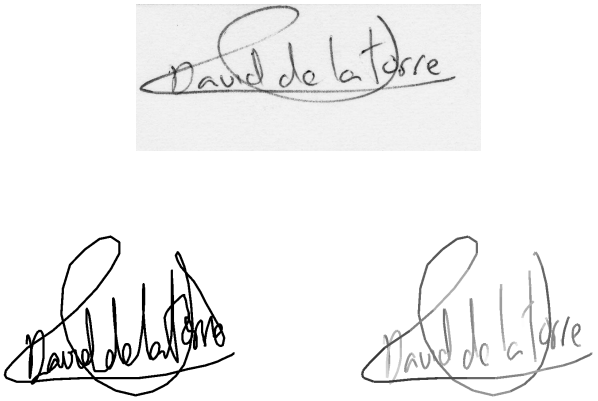

Figura 5.2: Comparación de la firma sintética con la firma original. Izquierda, versión 5. Derecha, versión 6

Entrenamiento de la CNN y creación de los modelos

Para cada una de las versiones se entrena la CNN con las 15 firmas genuinas de los 278 usuarios del conjunto de desarrollo como se ha expuesto en la Sección 3.2. y de ellos se extraen los 6 modelos diferentes.

Luego el conjunto de explotación de cada versión usará su respectivo modelo como extractor de características, ya que la intención es ver que versión funciona mejor con su modelo específico.

Entrenamiento de los clasificadores

El entrenamiento del SVM consiste, para cada usuario del conjunto de explotación, en introducir 10 firmas genuinas de éste como muestras positivas y 10 firmas genuinas del resto de usuarios, dando un total de 10x71 = 710 muestras negativas.

La división se lleva a cabo escogiendo aleatoriamente las 10 firmas para el entrenamiento y las 5 de test. Este proceso se repite 10 veces dando lugar a 10 resultados diferentes. Después se calcula la media y la desviación típica del EER Random (global threshold), EER Random (users thresholds), EER Skilled (global threshold) y EER Skilled (users thresholds).

5.2.3. Evaluación de los resultados

El primer paso es comparar los resultados obtenidos con el modelo SigNet como extractor de características a crear un modelo específico con la base de datos MCYT completa.

Como se puede apreciar en la Tabla 5.3 el modelo específico de cada versión obtiene mejores resultados que usar el modelo SigNet. Nos centramos solo en el **EER Skilled (Users thresholds)** ya que es usado para comparar con el estado del arte. Dentro de las diferentes versión, para el tamaño de píxel, la que mejores resultados obtiene es la versión 3, con un tamaño de píxel 3 y un resultado de **5.088 (+- 0.866)**. Este resultado aún está lejos del mejor obtenido por el estado del arte **2.87 (+- 0.42)** pero se utilizará como punto de decisión para el tamaño del píxel utilizado en las otras dos versiones.

Versión	Modelo	EER Skilled (global threshold)	EER Skilled (users thresholds)	EER Random (global threshold)	EER Random (users thresholds)
Versión 1	SigNet	10.519 (+- 0.669)	7.227 (+- 1.084)	0.356 (+- 0.194)	0.015 (+- 0.016)
	Versión 1	9.227 (+- 0.956)	5.426 (+- 0.826)	0.083 (+- 0.102)	0.004 (+- 0.009)
Versión 2	SigNet	12.782 (+- 0.969)	10.199 (+- 0.980)	1.221 (+- 0.328)	1.221 (+- 0.328)3
	Versión 2	8.704 (+- 0.984)	5.148 (+- 0.720)	0.160 (+- 0.158)	0.031 (+- 0.055)
Versión 3	SigNet	14.255 (+- 0.736)	10.931 (+- 1.143)	1.762 (+- 0.732)	0.642 (+- 0.359)
	Versión 3	8.940 (+- 0.937)	**5.088 (+- 0.866)**	0.071 (+- 0.093)	0.005 (+- 0.012)
Versión 4	SigNet	14.796 (+- 0.654)	12.167 (+- 0.726)	2.539 (+- 0.673)	0.920 (+- 0.499)
	Versión 4	9.060 (+- 0.754)	5.625 (+- 0.950)	0.087 (+- 0.104)	0.007 (+- 0.017)

Cuadro 5.3: Tabla de resultados para MCYT-72 sintética con diferentes versiones y comparando el uso del modelo SigNet con el específico de esa versión.

La Figura 5.3 expresa la evolución del tamaño del píxel frente al EER Skilled (Users Thresholds), se puede apreciar que el que mejor resultado obtiene es el tamaño 3 para el modelo específico de esa versión. Por otro lado, al utilizar el modelo SigNet se encuentran resultados peores. Aunque sea decreciente según disminuye el tamaño del píxel no tiene sentido llegar a un tamaño menor que uno ya la firma se vuelve prácticamente ilegible y se pierde la información esencial de cada usuario.

Una vez decidido el tamaño del píxel se prueba con las versiones 5 y 6. En la Tabla 5.4 se aprecia para la versión 6 (con información de presión) que alcanza un EER de **2.699 (+- 0.614)**, menor que el resultado del estado del arte mejor, 2.87 (+- 0.42). Por lo tanto, la versión 6 mejora el estado del arte analizado y va a ser la utilizada para los siguientes experimentos con otras bases de datos.

Es por esto que se puede concluir que los resultados son mejores cuando se utiliza un modelo específico para firma sintética en vez del modelo para firmas originales.

Más adelante se expondrán los resultados del mismo experimento pero usando un modelo global entrenado con un mayor número de usuarios y diferentes bases de datos que permita

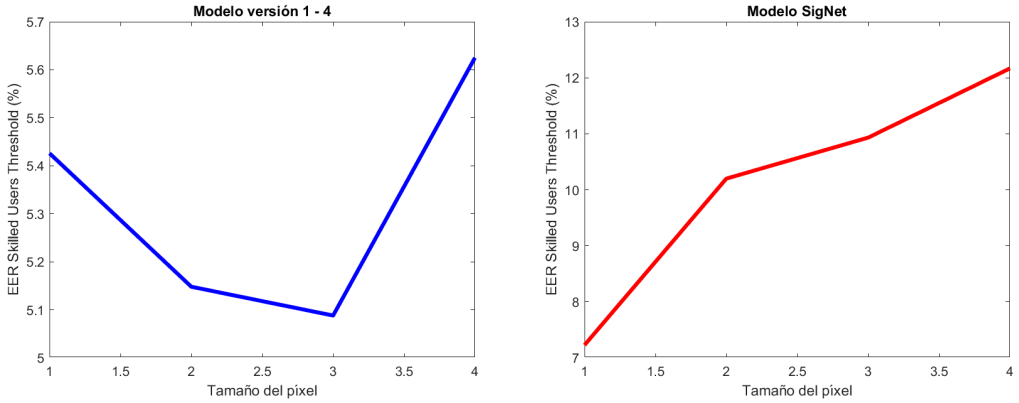

Figura 5.3: Comparación de utilizar diferentes tamaños de pixel en la extracción de la firma sintética con modelos específicos (versiones de 1-4) frente a utilizar el modelo SigNet.

Versión	Modelo	EER Skilled (global threshold)	EER Skilled (users thresholds)	EER Random (global threshold)	EER Random (users thresholds)
Versión 5	SigNet	12.519 (+- 0.954)	9.699 (+- 1.017)	1.512 (+- 0.575)	0.502 (+- 0.290)
	Versión 5	8.213 (+- 0.557)	4.463 (+- 0.735)	0.084 (+- 0.096)	0.001 (+- 0.001)
Versión 6	SigNet	6.019 (+- 0.407)	3.495 (+- 0.660)	0.246 (+- 0.205)	0.010 (+- 0.007)
	Versión 6	5.940 (+- 0.522)	**2.699 (+- 0.614)**	0.186 (+- 0.162)	0.017 (+- 0.043)

Cuadro 5.4: Tabla de resultados para MCYT-72 sintética con diferentes versiones y comparando el uso del modelo SigNet con el específico de esa versión.

mejorar aún más los resultados obtenidos hasta el momento. A su vez se probará a mejorar el rendimiento del SVM introduciendo también firmas del conjunto de desarrollo como muestras negativas y de este modo optimizar el rendimiento.

5.3. Creación de un modelo global

La finalidad de este experimento es crear un modelo global a partir de la base de datos definitiva descrita en la Sección 4.5. que se compone de MCYT, BioSecure DS2 y BioSecurID. A su vez se pondrá a prueba el sistema con diferente número de firmas genuinas y random, por usuario, para el entrenamiento del SVM.

Todo esto en un escenario más realista donde el conjunto de explotación no es utilizado como muestras negativas en los clasificadores, en este caso se utiliza el conjunto de desarrollo, dando por sentado que es más fácil para un sistema real contar con estos usuarios para entrenar el SVM que con los usuarios de explotación, los cuáles se quiere identificar.

Destacar que como en el modelo SigNet, en ningún momento se utilizan las falsificaciones de los usuarios para entrenar la red neuronal. En un futuro se podría hacer un estudio sobre como mejoraría el sistema con esa implementación.

5.3.1. Protocolo experimental

Para el entrenamiento y creación del modelo se utilizará la división descrita en la Tabla 5.5.

Base de Datos	**#Usuarios**	**#firmas**	**Total (#firmas)**
MCYT	230	16	230x16 = 3680
BioSecure DS2	510		510x16 = 8160
BioSecurID	268		268x16 = 4288
Total	1008		1008x16 = **16128**

Cuadro 5.5: Número de firmas utilizadas para el entrenamiento de la CNN.

Para entrenar los clasificadores se toma un escenario más realista como se ha descrito anteriormente. De este modo se entrenarán con las firmas genuinas del usuario como muestras positivas y las firmas genuinas de los 1008 usuarios del conjunto de desarrollo como muestras negativas. Por lo tanto, el sistema no se habrá familiarizado con el conjunto de explotación hasta el momento de la evaluación.

Además, la división de las firmas genuinas del usuario para entrenamiento y test no se realizará de manera aleatoria, sino en orden de escritura. Es decir, se utilizarán las firmas de las sesiones más recientes para el testeo y las de las más antiguas para el entrenamiento, replicando un sistema real ya que las bases de datos tienen las firmas ordenas cronológicamente.

- **Experimento 1**: se utilizan las 12 primeras firmas genuinas en la fase de entrenamiento y las 4 últimas en la fase de test.
- **Experimento 2**: se utilizan las 8 primeras firmas genuinas en la fase de entrenamiento y las 8 últimas en la fase de test.
- **Experimento 3**: se utilizan las 4 primeras firmas genuinas en la fase de entrenamiento y las 12 últimas en la fase de test.

En la Tabla 5.6 se describe el entrenamiento y test de los clasificadores llevado a cabo en este experimento.

Base de Datos	#Usuarios	Experimento	Entrenamiento		Test		
			Genuinas	Random	Genuinas	Random	Skilled
MCYT	100	1	12	16 x 1008 = 16128	4	4 x 99 = 366	12
		2	8	16 x 1008 = 16128	8	8 x 99 = 792	12
		3	4	16 x 1008 = 16128	12	12 x 99 = 1188	12
BioSecure DS2	140	1	12	16 x 1008 = 16128	4	4 x 139 = 556	12
		2	8	16 x 1008 = 16128	8	8 x 139 = 1112	12
		3	4	16 x 1008 = 16128	12	12 x 139 = 1668	12
BioSecurID	132	1	12	16 x 1008 = 16128	4	4 x 131 = 524	12
		2	8	16 x 1008 = 16128	8	8 x 131 = 1048	12
		3	4	16 x 1008 = 16128	12	12 x 131 = 1572	12

Cuadro 5.6: Protocolo experimental de los clasificadores SVM para los diferentes experimentos según cada base de datos

A su vez se realizará un experimento extra en el que se utilizará solamente el 50 % de los usuarios del conjunto de desarrollo como muestras negativas en la fase de entrenamiento. Es decir, 504 usuarios en vez de los 1008. De este modo se podrá ver la pérdida de calidad al disminuir las muestras negativas del clasificador.

5.3.2. Desarrollo experimental

Debido a los resultados de la Sección 5.2. se decide crear la firma sintética utilizando la **Versión 6** (tamaño de píxel 3, e información de presión), ya que es la versión que mejor resultados ha dado para MCYT-72. Una vez extraídas las imágenes, la red neuronal es entrenada con el conjunto de desarrollo de las tres bases de datos que cuenta con 16 firmas genuinas por usuario. Todo el proceso de entrenamiento viene descrito en la Sección 3.2. de este trabajo.

El modelo creado a partir de este desarrollo se denominará **modelo global** ya que engloba varias bases de datos diferentes, esperando así que extraiga unas características más robustas que permitan a los clasificadores diferenciar y verificar la identidad del usuario de manera más óptima.

Tras la creación del modelo se extraen las 2048 características del conjunto de explotación que serán utilizadas por los clasificadores. Éstos son entrenados con según el protocolo experimental. Ya que la división de las firmas genuinas no es aleatoria, solo se repite el proceso una vez.

5.3.3. Evaluación de los resultados

Los resultados se pueden apreciar en la Tabla 5.7. Cuanto menor es el número de firmas genuinas utilizadas en el entrenamiento, peores son los resultados para cualquier EER. A su vez, la base de datos BioSecure DS2 es la que peores resultados obtiene, esto es sorprendente si se tiene en cuenta que es la base de datos que mayor número de firmas introduce en el entrenamiento de la CNN.

Base de Datos	Experimento	EER Random		EER Skilled	
		Global Threshold	Users Thresholds	Global Threshold	Users Thresholds
MCYT	1	0.210	0.009	6.625	3.708
	2	0.307	0.028	7.458	3.833
	3	0.951	0.134	10.000	7.750
BioSecure DS2	1	3.234	0.847	13.333	7.560
	2	2.201	0.918	11.027	7.485
	3	2.161	1.196	12.440	8.244
BioSecurID	1	0.599	0.272	8.617	3.598
	2	1.449	0.494	10.906	8.002
	3	2.129	0.494	13.542	10.669

Cuadro 5.7: Resultados para los diferentes experimentos según las bases de datos.

Cabe destacar que el comportamiento al disminuir el número de firmas es dependiente de la base de datos. Esto puede ser debido a que no se realiza una división aleatoria, por lo tanto habrá firmas de peor calidad que cuando pasan de estar en entrenamiento a test cambian drásticamente los resultados.

En la Fig. 5.4 se puede apreciar una comparación de las tres bases de datos en EER Skilled (Users Thresholds) variando el número de firmas genuinas por usuario en el entrenamiento del clasificador. La línea continua muestra el resultado de utilizar todo el conjunto de desarrollo como muestras negativas, y la discontinua de utilizar solamente el 50 %, es decir 504 usuarios.

No hay una pérdida significativa si se tiene en cuenta que se ha disminuido a la mitad el número de usuarios utilizados como muestras negativas reduciendo el coste computacional del sistema.

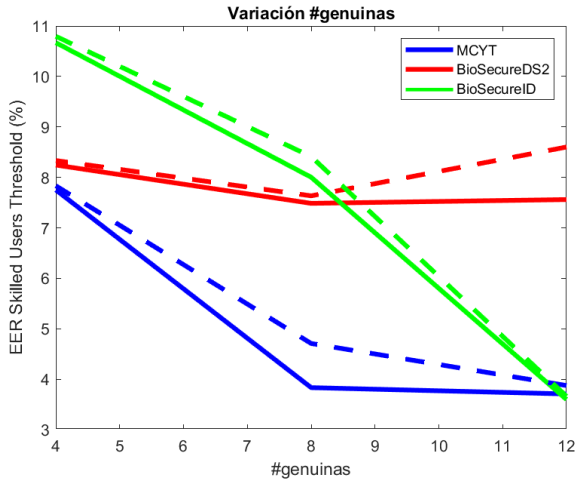

Figura 5.4: Comportamiento de las diferentes BBDD. Línea continua, utilizando el 100 % del conjunto de desarrollo. Línea discontinua, utilizando el 50 %.

Por último, cabe recordar que las bases de datos originales MCYT y BioSecure DS2 cuentan con un número bastante mayor tanto de firmas genuinas por usuario (MCYT cuenta con 25 firmas genuinas y BioSecure DS2 con 30), por lo que se podrían conseguir mejores resultados utilizando el modelo extractor de características global pero aumentando el número de firmas genuinas en el entrenamiento de los clasificadores. En el protocolo experimental no se tuvo en cuenta para igualar las tres bases de datos al mismo número de firmas genuinas.

5.4. Experimentación con los clasificadores

5.4.1. Protocolo y desarrollo experimental

El protocolo y desarrollo experimental es el mismo que en la Sección 5.2. salvo que se usará el **modelo global** creado en la Sección 5.3. Y se harán variaciones en el número de muestras negativas de los clasificadores. Por lo tanto, se hará uso de la versión 6 de MCYT-72.

Se centra en la diferencia que hay entre usar los usuarios del conjunto de desarrollo y/o explotación. Es decir, las muestras negativas que ayudan al SVM a crear el hiperplano. Se tratan tres casos posibles:

- **Caso 1**: se utilizan solamente los usuarios del conjunto de explotación, 71 usuarios por SVM como muestras negativas. Igual que en la Sección 5.2. no se introducen las firmas del usuario que serán utilizadas en la fase de test.
- **Caso 2**: se utilizan solamente los usuarios del conjunto de desarrollo siendo un caso más realista ya que en principio un sistema no cuenta con los usuarios de explotación para entrenar los clasificadores, 258 usuarios por SVM como muestras negativas con todas las firmas genuinas, ya que los usuarios de desarrollo no se utilizan en la fase de test.
- **Caso 3**: se utilizan tanto los usuarios del conjunto de desarrollo como los de explotación, 329 usuarios por SVM como muestras negativas.

5.4.2. Evaluación de los resultados

En la Tabla 5.9 se puede apreciar los resultados, siendo desarrollo el conjunto de desarrollo y explotación el conjunto de explotación utilizados como muestras negativas en el entrenamiento del clasificador. Por lo tanto, la primera fila se corresponde con el caso 1 descrito anteriormente y así sucesivamente.

MCYT-72	**Caso**	**Usuarios**		**EER Skilled** (user thresholds)	**EER Random** (users thresholds)
		Desarrollo	**Explotación**		
Versión 6	1	0	71	2.199 (+- 0.577)	0.001 (+- 0.001)
	2	258	0	2.472 (+- 0.771)	0.012 (+- 0.010)
	3	258	71	**1.921 (+- 0.573)**	0.005 (+- 0.007)

Cuadro 5.8: Número de usuarios y del conjunto que provienen introducidos como muestras negativas en el clasificador y su correspondiente EER, tanto Skilled (user thresholds) como Random (users thresholds). Características extraídas con el modelo global.

Comparando el EER Random se aprecia que el mejor resultado es utilizando solamente el conjunto de explotación, mejor incluso que utilizando los dos conjuntos a la vez. Esto es debido a que en la fase de test se compara con los usuarios del conjunto de explotación, por lo tanto el sistema se familiariza con estos usuarios y le es más sencillo clasificarlos.

Comparando el EER Skilled. Al utilizar los usuarios del conjunto de explotación se obtienen mejores resultados que utilizando los del conjunto de desarrollo, aunque éstos últimos sean un número mayor de usuarios. Cosa que es sorprendente si se tiene en cuenta que el EER de skilled se extrae solamente comparando las firmas genuinas y falsificaciones del mismo usuario. A su vez, utilizar ambos conjuntos mejora aún más el sistema, haciéndolo más preciso.

5.5. Comparación con el estado del arte

Como se puede apreciar en la Tabla 5.9 se mejora el estado del arte analizado con el sistema propuesto. Esto puede ser debido a la utilización de la información de presión a la hora de extraer la firma sintética.

Referencia	**# Muestras**	**Características**	**EER**
Gilperez et al.[26]	10	Contours (chi squared distance)	6.44
Vargas et al.[28]	10	LBP (SVM)	7.08
Ooi et al.[29]	10	DRT + PCA (PNN)	9.87
Soleimani et al.[30]	10	HOG (DMML)	9.86
Sabourin et al.[1]	10	SigNet (SVM)	2.87 (+- 0.42)
Propuesto	**10**	**Global (SVM)**	**2.20 (+- 0.58)**

Cuadro 5.9: Comparación con el estado del arte en MCYT. EER de skilled forgeries.

Cabe destacar que los resultados no son directamente comparables ya que en este trabajo se utiliza firma sintética y no original. A su vez, debido a la falta de 3 usuarios de MCYT-75 para la creación de las firmas sintética, en este caso son 72 usuarios.

6

Conclusiones y trabajo futuro

Podemos concluir que los resultados han sido muy satisfactorios ya que se ha mejorado el estado del arte para MCYT-75, y al aplicarlo a otras bases de datos se siguen obteniendo buenos resultados, tanto en skilled forgeries como en random forgeries.

Se ha creado con éxito un sistema de verificación de firma estática sintética a través de CNNs que obtiene buenos resultados tanto contra random forgeries como skilled forgeries.

Cabe destacar la importancia que ha tenido la extracción de la firma off-line sintética, ya que al usar la información de los pen-ups y la presión el sistema mejoraba con creces. También, el entrenamiento de los clasificadores ha sido crucial, concluyendo que cuantas más muestras positivas y negativas se dispongan, mejor podrá verificar y autenticar a los usuarios.

La firma off-line no está tan lejos de la on-line aún de no contar con la información temporal. Además, debido a su gran facilidad de implementación y aceptabilidad por el usuario, en un futuro podría convertirse en uno de los sistemas biométricos más importantes.

A pesar de los buenos resultados obtenidos, existe un importante trabajo por hacer con el sistema. Algunos estudios que se podrían llevar a cabo en un futuro son los siguientes:

- Proponer distintos modelos CNN adaptados a las complejidades de los usuarios.
- Proponer el uso de otras arquitecturas más robustas que consideren distintos escenarios, como el uso del dedo.
- Proponer un sistema que tenga en cuenta tanto la firma off-line como la on-line, realizando una unión de ambos para mejorar los clasificadores.
- Utilizar modelos pre-entrenados en otros escenarios con mayores volúmenes de datos y adaptarlo al problema de firma (AlexNet, VGG19, ResNet, etc.).
- Cambiar la estructura de la CNN.
- Utilizar otro tipo de clasificadores.
- Realizar un estudio con clasificadores que no sean writer-dependent.
- Entrenar la red con *skilled forgeries* o con una mezcla de *skilled* y *random*.

Glosario de acrónimos

- **BiDA MDI-Sign Database**: *BiDA Lab Multiple Devices and Input Online Signature Database*
- **CNN**: *Convolutional Neural Networks*
- **DCT**: *Discret Cosine Transform*
- **DET**: *Detection Error Trade-off*
- **DTW**: *Dynamic Time Warping*
- **EER**: *Equal Error Rate*
- **FA**: Falsa Aceptación
- **FAR**: *False Acceptance Rate*
- **FR**: Falso Rechazo
- **FRR**: *False Rejection Rate*
- **HMM**: *Hidden Markov Models*
- **LSTM**: *Long Short-Term Memory*
- **NN**: *Neural Networks*
- **RNN**: *Recurrent Neural Networks*
- **SVM**: *Support Vector Machines*
- **WACOM**: *marca de dispositivos de captura específicos para la recogida de firmas*
- **WD**: *Writer Dependent*

Bibliografía

[1] Robert Sabourin Luiz G. Hafemann and Luiz S. Oliveira. Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70:163–176, 2017.

[2] Pablo Lázaro Herrasti. Recopilación y uso de datos masivos en sistemas de verificación de firma manuscrita estática. *Universidad Autónoma de Madrid*, 2018.

[3] J. Fierrez, A. Pozo, M. Martinez-Diaz, J. Galbally, and A. Morales. Benchmarking touchscreen biometrics for mobile authentication. *IEEE Trans. on Information Forensics and Security*, 13(11):2720–2733, November 2018.

[4] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Incorporating touch biometrics to mobile one-time passwords: Exploration of digits. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPR-W*, June 2018.

[5] La Redacción. Biometría en gobierno: 4 casos exitosos en asia. *Tecnología en Gobierno*, 2018.

[6] R. Vera-Rodriguez, J. Fierrez, J. Ortega-Garcia, A. Acien, and R. Tolosana. e-biosign tool: Towards scientific assessment of dynamic signatures under forensic conditions. In *Proc. IEEE BTAS*, September 2015.

[7] A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of multibiometrics: human recognition systems; 1.* Springer, Dordrecht, 2006.

[8] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[9] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

[10] G. E. Hinton. Boltzmann machine. *Scholarpedia*, 2(5):1668, 2007.

[11] S. Lawrence, C.L. Giles, Ah Chung Tsoi, and A.D. Back. Face recognition: a convolutional neural-network approach. *Neural Networks, IEEE Transactions on*, 8(1):98–113, January 1997.

[12] O. Costilla Reyes, R. Vera-Rodriguez, P. Scully, and K. B. Ozanyan. Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (99), 2018.

[13] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber. Convolutional neural network committees for handwritten character classification. In *2011 International Conference on Document Analysis and Recognition*, pages 1135–1139, 2011.

[14] A. Graves and N. Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In *Proc. ICML*, pages 1764–1772. JMLR.org, 2014.

[15] A. Petrosian, D. Prokhorov, R. Homan, R. Dasheiff, and D. Wunsch. Recurrent neural network based prediction of epileptic seizures in intra- and extracranial eeg. *Neurocomputing*, 30(1):201 – 218, 2000.

[16] R. Tolosana, R. Vera-Rodriguez, R. Guest, J. Fierrez, and J. Ortega-Garcia. Complexity-based biometric signature verification. In *14th IAPR International Conference on Document Analysis and Recognition*, 2017.

[17] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally. Mobile signature verification: Feature robustness and performance comparison. *IET Biometrics*, 3(4):267–277, December 2014.

[18] N. Houmani, S. Garcia-Salicetti, B. Dorizzi, and *et al.* Biosecure signature evaluation campaign (esra'2011): evaluating systems on quality-based categories of skilled forgeries. In *Proc. IJCB*, pages 1–10, Oct 2011.

[19] M. Martinez-Diaz and J. Fierrez. *Signature Databases and Evaluation*, pages 1367–1375. Springer, 2015. ISBN 978-1-4899-7487-7, re-edited from 2009.

[20] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez. Hmm-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, 28(16):2325–2334, December 2007.

[21] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia. Benchmarking desktop and mobile handwriting across cots devices: the e-biosign biometric database. *PLOS ONE*, 5(12), 2017.

[22] A. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recogn.*, 38(12):2270–2285, December 2005.

[23] Keiron O'Shea and Ryan Nash. An introduction to convolutional neural networks. 2015.

[24] C. Travieso J. Alonso J. Vargas, M. Ferrer. Off-line handwritten signature gpds960. *Document Analysis and Recognition, 9th International Conference*, page 764–768, 2007.

[25] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez, V. Espinosa, A. Satue, I. Hernaez, J. J. Igarza, C. Vivaracho, D. Escudero, and Q. I. Moro. Mcyt baseline corpus: A bimodal biometric database. *IEE Proceedings Vision, Image and Signal Processing, Special Issue on Biometrics on the Internet*, 150(6):395–401, December 2003.

[26] S. Pecharroman J. Fierrez J. Ortega-Garcia A. Gilperez, F. Alonso-Fernandez. Off-line signature verification using contour features. *11th International Conference on Frontiers in Handwriting Recognition, Montreal, Quebec–Canada*, 2008.

[27] Y.Y. Tang T. Zhang J. Wen, B. Fang. Model-based signature verification with rotation invariant features. *Pattern Recognit.*, 42(7):1458–1466, 2009.

[28] C.M. Travieso J.B. Alonso J.F. Vargas, M.A. Ferrer. Off-line signature verification based on grey level information using texture features. *Pattern Recognit.*, 44(2):375–385, 2011.

[29] Y.H. Pang B.Y. Hiew S.Y. Ooi, A.B.J. Teoh. Image-based handwritten signature verification using hybrid methods of discrete radon transform. *Principal component analysis and probabilistic neural network*, 40(2):274–282, 2016.

[30] K. Fouladi A. Soleimani, B.N. Araabi. Deep multitask metric learning for offline signature verification. *Pattern Recognit.*, page 84–90, 2016.

[31] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. R. Freire, F. Alonso-Fernandez, and D. Ramos *et al.* Biosecurid: A multimodal biometric database. *Pattern Analysis and Applications*, 13(2):235–246, May 2010.

[32] J. Ortega-Garcia, J. Fierrez, F. Alonso-Fernandez, J. Galbally, and *et al.* The multi-scenario multi-environment biosecure multimodal database (bmdb). *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(6):1097–1111, June 2010.